



## **XVII Jornadas Expania**

### **Taller de Alma Analytics y Visualización de Datos**

**Madrid, 15/12/2020**

**Mercedes Echeverría**



# Taller de Analytics y Visualización de Datos

04 Dec 2020

## ¡Bienvenido al Taller de Alma-Analytics y Visualización de datos de Expanla 2020!

El objetivo del 'Taller de Alma-Analytics y Visualización de Datos' es presentar los nuevos desarrollos de estas herramientas en función de las actualizaciones introducidas por Exlibris en Octubre de 2020.

La visión del Taller es exponer contenidos teóricos y casos prácticos que puedan servir para obtener información relevante de Alma, de forma que ayude especialmente para adoptar decisiones fundamentadas en datos e impulse procesos eficientes.

El espíritu del Taller sigue siendo una actualización constante de los contenidos de Alma-Analytics dentro de un contexto de colaboración y de intercambio de información entre los participantes. Para lograr este propósito, dentro de las limitaciones actuales, contamos con el programa Confluence (Atlassian) que nos va a permitir mantener un diálogo directo y la plataforma Zoom que facilitará la presentación por videoconferencia.

### AGENDA

**Fecha 15/12/20. Horario: 9:30 – 13:30**

#### Alma – Analytics (OAS)

De 9:30 – 11:30 h.

1. Informes con utilización de funciones SQL
2. Informes con filtros avanzados (Petición de datos de Panel de Control)
3. Configurar un Portal de Analytics
4. Objetos de Analytics

*Preguntas:* 11:15 – 11:30 h.

#### Alma - Visualización de Datos

De 11:45 :-13:30 h. h.

1. Proyecto: Clúster
2. Proyecto: ARIMA
3. Combinación de datos de diferentes áreas

*Preguntas:* 13:15 – 13:30 h.

### Contenidos teóricos del Taller de Analytics

1. Alma Analytics (OAS)
2. Visualización de datos (DV)
3. Análisis estadísticos

## Indice

Taller de Analytics y Visualización de Datos .....	1
1.1 Alma Analytics (OAS).....	4
1.1.1 Introducción.....	5
1.1.2 Navegación en Analytics.....	6
1.1.3 Creación de informes.....	11
1.1.4 Filtros.....	27
1.1.5 Paneles de control (Dashboards).....	35
1.1.6 Portal de Analytics.....	41
1.1.7 Objetos de Analytics.....	50
1.2. Visualización de Datos (DV).....	53
1.2.1 Introducción.....	54
1.2.2 Proyectos.....	59
1.2.2.1 Proyecto 1.....	60
1.2.2.2 Proyecto 2.....	66
1.2.3 Creación de proyectos con datos de diferentes áreas.....	71
1.3. Análisis estadísticos.....	79
1.3.1 Análisis de clúster.....	80
1.3.2 Series temporales y modelos ARIMA.....	90
1.3.3 Regresión lineal simple y línea de tendencia.....	101

# Alma Analytics (OAS)

## Introducción

1. [Navegación en Analytics](#)
2. [Creación de informes](#)
3. [Filtros](#)
4. [Paneles de control \(Dashboards\)](#)
5. [Portal de Analytics](#)
6. [Objetos de Analytics](#)



## Introducción

Alma Analytics es el módulo de estadísticas de Alma que realiza informes basados en las actividades que llevan a cabo las bibliotecas. Los informes generados ayudan a comprender el funcionamiento de los servicios de la biblioteca, y sirven especialmente para tomar decisiones fundamentadas en datos que faciliten la gestión de la biblioteca e impulsen procesos eficientes.

Analytics es un módulo separado de Alma con el que se sincroniza cada 24 horas. La actualización de los informes de Analytics se produce a partir de los datos generados por Alma relacionados con los procesos desarrollados el día anterior. Este módulo facilita la creación de informes originales, permite la utilización de informes predefinidos por ExLibris, la reutilización de informes generados por bibliotecas usuarias de Alma, y la elaboración de informes creados por la propia institución, bien para ser compartidos internamente, o externamente por bibliotecas de otras instituciones.

La versión actual de Analytics (Octubre 2020), está construida sobre Oracle (OBIEE 12) y utiliza el lenguaje de consulta estructurada SQL (Structure Query Language). El portfolio de Oracle respecto Analytics está formado por 'Oracle Analytics Server' (OAS) y la herramienta independiente 'Data visualization' (DV). La tecnología de OAS se sustenta en la analítica aumentada, cuyas bases son la inteligencia artificial (IA), machine learning (ML) y el procesamiento del lenguaje natural (NLP). Ofrece ventajas muy precisas para consultas en lenguaje natural, análisis avanzado automatizado, extracción de datos visuales de forma automatizada y análisis estadísticos inferenciales.

Analytics facilita datos actualizados de los procesos bibliotecarios, ofrece formatos fácilmente legibles, incluyendo tablas y gráficos, y proporciona una correcta entrega de la información a los bibliotecarios de la institución.

Esta Guía describe los componentes esenciales de la nueva versión de Analytics, y expone las actividades básicas para la creación de informes basados en datos procedentes de Alma.

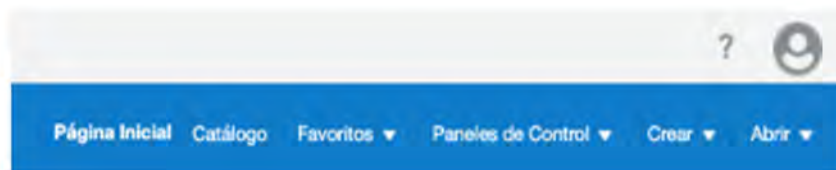
## 1. Navegación en Analytics

- 1. Perfil del usuario
- 2. Ayuda
- 3. Página inicial
- 4. Catálogo
- 5. Favoritos
- 5. Paneles de control o dashboards
- 6. Crear
- 7. Abrir

En la parte superior se encuentra el menú principal de Analytics, diferenciado en dos filas :

--> Fila superior: *Perfil del usuario, Ayuda*

--> Fila inferior: *Páginas inicial, Catálogo, Favoritos, Paneles de Control, Crear, Abrir*





## 1. Perfil del usuario

Para crear informes es preciso tener asignado el perfil 'Design Analytics Role'.

Alma define dos tipos de usuarios de Analytics, aquellos que pueden realizar informes (*report designers*) y quienes sólo pueden visualizar los informes (*report consumers*).

## 2. Ayuda

Ayuda de Oracle Analytics Server (OAS)

## 3. Página inicial

Contiene dos secciones diferenciadas:

-En la parte izquierda se encuentran los apartados:

- *Crear*: el contenido de este apartado contiene herramientas para la generación de informes
- *Informes publicados*: no accesible
- *Inteligencia procesable*: funcionalidades externas.
- *Trabajos*: no accesible

-El lienzo central contiene dos apartados:

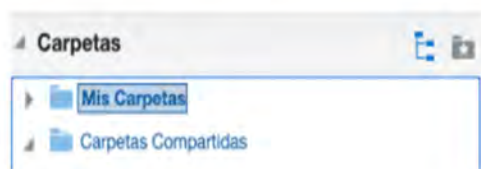
- *Reciente*: muestra los informes en los que ha estado trabajando recientemente.
- *Otros*: muestra los informes más frecuentemente consultados por los usuarios



#### 4. Catálogo

Muestra el directorio completo de Analytics al que tiene acceso el diseñador.

El directorio contiene dos subdivisiones principales:



→ **Mis carpetas:** es el área privada, visible sólo para el propietario de la cuenta de Analytics.

→ **Carpets compartidas:**

Contiene informes compartidos y creados por otras bibliotecas de Alma. Es una buena fuente de inspiración e ideas, pudiendo ser reutilizados.

✓ Tenga en cuenta:

Cuando reutilice informes de otras universidades, **se recomienda archivarlos en sus carpetas locales para respetar el fichero de origen.**

Esta sección consta de cuatro apartados:





a) **Alma:** contiene informes prediseñados por ExLibris. Esta carpeta contiene subcarpetas con plantillas de informes definidos para todas las áreas temáticas de Analytics.

b) **Community:** contiene informes de bibliotecas miembros de Alma, puestos a disposición de los usuarios de Alma-Analytics para que puedan ser compartidos y reutilizados.

Esta sección se subdivide en:

- Analytics Master Class:* contiene informes prediseñados sobre gestión mantenimiento de colecciones. Esta sección está en fase de elaboración.
- Consortia:* informes procedentes de consorcios de bibliotecas.
- Ex Libris Development:* contiene informes desarrollados por ExLibris sobre comparación de colecciones, DARA, colecciones electrónicas y los creados por Yoel Kortick, Senior Librarian de Ex Libris sobre mantenimiento de colecciones.
- Institutions:* informes y dashboards compartidos entre bibliotecas miembros de Alma.
- Regional User Group:* carpeta de informes compartidos de grupos de usuarios a nivel territorial.
- Shared reports:* contiene subcarpetas de áreas temáticas preestablecidas por ExLibris donde cualquier institución puede contribuir con sus propios informes.
- Support:* cualquier institución puede colocar informes con casos en esta subcarpeta y notificarlo a ExLibris.

c) **Leganto:** aplicación de Ex Libris destinada a gestionar la bibliografía recomendada.

d) **Espacio compartido para cada universidad**

Este espacio es visible sólo para los bibliotecarios de la propia institución.

## 5. Paneles de control o dashboards

Colección de informes reunidos bajo un área común de visualización.



## 6. Crear

Lista de objetos que pueden ser creados: *análisis, panel de control, filtro, petición de datos de panel de control, condición.*



## 7. Abrir

Permite seleccionar un objeto con el que se quiere trabajar. En la primera columna se encuentran los objetos más recientemente creados, vistos o modificados. En la segunda columna, se encontrarían los objetos más solicitados.



## 2. Creación de informes

- Introducción
- 1. Areas temáticas
  - Descripción de un área temática
  - Criterios
- Filtros de selección
- Formato de los informes
  - → Estilo del informe
  - → Formato columna
  - → Formato de los datos.
  - → Formato condicional
  - → Interacción
- Añadir un gráfico a un informe
- Representaciones gráficas
  - → Diagrama de barras
  - → Diagrama de sectores
  - → Histogramas
  - → Series temporales
  - → Nube de puntos
- Materiales de apoyo

## Introducción

Para crear un análisis vaya al menú superior de Analytics y seleccione:

- a) **Crear**
- b) Pulse en **Análisis**
- c) Precise un '**Area temática**'

### 1. Areas temáticas

Áreas temáticas	Contenido
AP Usage	Monitoriza el uso de APPs
Analytics objects	Ofrece datos sobre la lista objetos, widgets e informes programados en Analytics.
Analytics Usage Tracking	Facilita el seguimiento del uso de Analytics con detalles de las consultas, sobre las fechas de las consultas y áreas temáticas más usadas.
Benchmark	Facilita la comparación de indicadores de rendimiento entre instituciones similares a través KPIs (Key Performance Indicators) de inventarios, adquisiciones, préstamos, reservas, informes COUNTER.
Borrowing Requests	Información sobre las peticiones de demanda de préstamo interbibliotecario.
Course Reserves	Reservas bibliográficas para cursos. No aplica
Digital Inventory	Información sobre el inventario digital: número de ficheros digitales, objetos digitales, número de descargas, así como detalles de los ficheros, objetos.
Digital Usage	Número de visitas, descargas o fallos de acceso a los objetos digitales, así como detalles de los objetos o fechas de consulta.
E-Inventory	Informes y dashboards de inventarios de portfolios y colecciones electrónicas.
Events	Es un área genérica que investiga eventos generados en Alma. Por ejemplo, presenta actividades de catalogación como la creación de registros, borrado y actualización de MMS-IDs.
Fines and fees	Información sobre sanciones por retrasos o por otras políticas de sanción que aplique la biblioteca. Ofrece una lista de usuarios multados
Fulfillment	Informes y dashboards sobre el área de préstamo.
Funds Expenditure	Información económica sobre fondos asignados, gasto, impuestos aplicados. Detalla información sobre los impuestos aplicados.



<a href="#">Leganto</a>	Es la aplicación de ExLbris orientada a gestionar las listas de bibliografía recomendada. Esta funcionalidad no ha sido contratada por el Consorcio Madroño.
<a href="#">Lending Requets</a>	Información sobre el suministro de préstamo interbibliotecario a otras bibliotecas.
<a href="#">Licenses</a>	Información sobre las licencias de los recursos electrónicos. Responde a preguntas como el número de portfolios y recursos por licencia y los gastos totales por licencias y líneas de pedidos para años fiscales determinados.
<a href="#">Link Resolver Usage</a>	Ofrece datos sobre el resolvidor de enlaces OpenURL y una amplia perspectiva sobre cómo funciona este estándar en Alma. Esta información es útil para analizar el desarrollo de la colección electrónica. Ofrece datos de usabilidad de los recursos electrónicos.
<a href="#">Physical items</a>	Ofrece una información detallada del inventario de fondos en papel: títulos, localizaciones, holdings, última fecha de préstamo y transacciones por compra.
<a href="#">Physical Items Historical Events</a>	Ofrece información de los eventos históricos de los ítems físicos y puede responder a cuestiones como el tránsito de los ítems, localizaciones temporales o cambio del estatus de los ítems.
<a href="#">Purchase Requests</a>	Informe que recopila las solicitudes de compra de los usuarios que usan Primo para que las bibliotecas compren un recurso.
<a href="#">Requests</a>	Suministra información de las reservas de los ejemplares físicos: reservas activas, expiradas, solicitudes de reservas que tienen más demanda, reservas por patrón de préstamo o por tipo de material.
<a href="#">Titles</a>	Ofrece información sobre el número de títulos de ítems físicos, holdings, portfolios, número de colecciones electrónicas. Suministra medidas del uso de los recursos bibliográficos o cómo vincular los holdings físicos con los portfolios para disponer de una representación de la colección.
<a href="#">Usage Data</a>	Suministra datos de los informes COUNTER: (Release 5): DR (Database Reports), PR (Platform Reports), TR-B (Title Master Requests -Books) TR-J (Title Master Request-Journals).
<a href="#">Users</a>	Facilita información detallada de los usuarios como roles, categorías, tipologías, grupos de usuarios y datos demográficos.

### Descripción de un área temática

Cuando seleccione un área temática se abrirá el espacio de trabajo. Los elementos fundamentales son:

#### a) Menú superior

Criterios de selección del análisis  
Resultados del análisis  
Peticiones de datos, usado para valores que no han sido prefiltrados  
Avanzado: análisis XML y sintaxis SQL  
Guardar: actualizar los datos guardados  
Guardar como: salvar el informe

#### b) Columna izquierda.

Contiene los criterios de selección, integrados por datos cuantitativos y descriptivos.

#### c) Lienzo central

Dividido en dos áreas, 'criterios seleccionados' y 'filtros'.



### Criterios

Los tipos de criterios pueden ser:

→ **Datos cuantitativos:** valores de medidas. Se identifican por el estilo de la columna, color amarillo.

- Num of Loans (In House + Not In House)
- Num of Loans (In House)
- Num of Loans (Not In House)

→ **Atributos:** valores categóricos, campos descriptivos textuales, o números discretos que se comportan como texto. Se identifican por el estilo de la columna, color azul.

- Material Type
- Modification Date
- Modified By

→ **Columnas jerárquicas:** valores ordenados por niveles o por criterios padre-hijo.

Por ejemplo, 'Location'.

Location

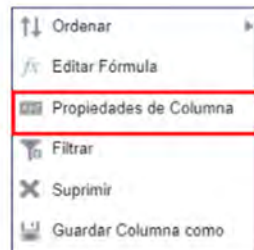
### Filtros de selección

Consulte la información detallada en el apartado 'Filtros'.



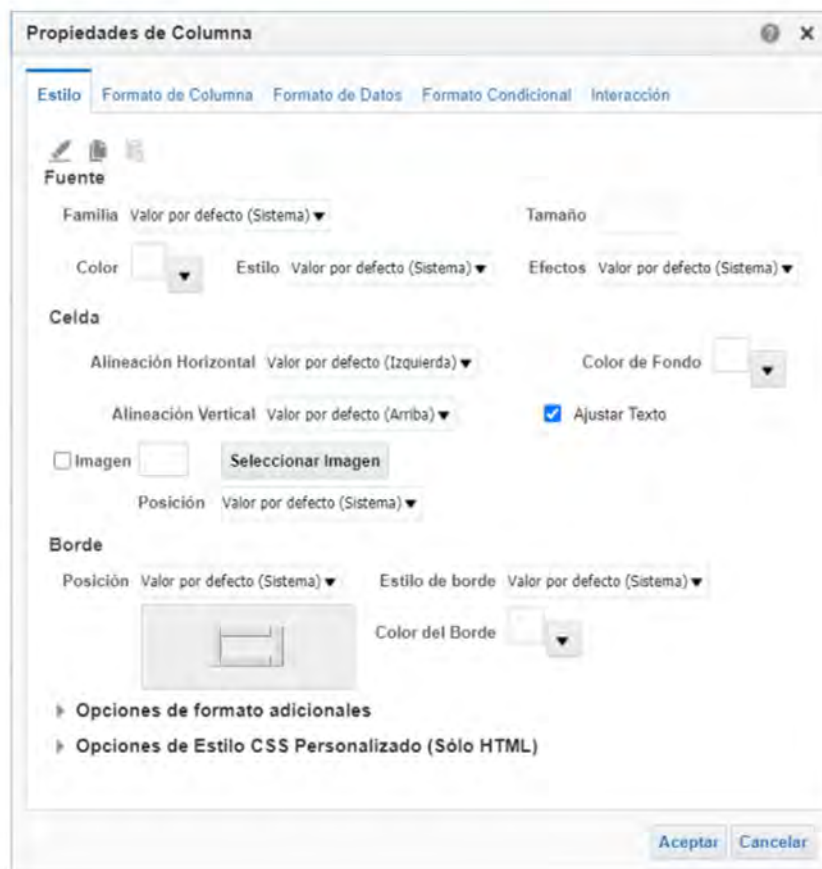
## Formato de los informes

Los criterios de los informes disponen de una tabla de propiedades que puede ser configurable por: *estilo, formato de la columna, formato de los datos, formato condicional e interacción*



### → Estilo del informe

El estilo del informe puede modificarse a partir de las fuentes, estilo de las celdas y bordes.



Fund. Fiscal Year	Fund Ledger Name	Location Name	Permanent Call Number	Barcode	Title	Editor	Lugar de Publicación	Publication Date	Tipo de material Registro ejemplar	Tipo de material Registro bibliográfico	Revisión Tipo (Bibliographic details)	Item Creation Month	Fecha de creación Registro ejemplar	Fecha de recepción Registro ejemplar	Fecha de creación Registro bibliográfico	Política del ítem
FY. 2018	MONOGRAFÍAS	En proceso	PREDICIONES	42872	CLU: cuadernos de literatura infantil y juvenil	Torre de Papel	Barcelona	1988	Journal	Journal	Journal - Physical	March	13/03/2025		15/03/2018	None
				42854	Padrón Periódico de Rousa y dificultad Rousa	SEPTI	Madrid	1969	Journal	Journal	Journal - Physical	March	13/03/2025		15/03/2018	None

#### → Formato columna

Aplica el estilo de las cabeceras de las columnas.

--> *Título de la cabecera*, puede ser modificado y personalizado.

**Propiedades de Columna**

Estilo **Formato de Columna** Formato de Datos Formato Condicional Interacción

**Cabeceras** ☐ Ocultar

Cabecera de Carpeta: Physical Item Details

Cabecera de Columna: Procedencia

☒ Cabeceras Personalizadas  
☐ Contiene marcador HTML

**Supresión de Valor**

☒ Suprimir ☐ Repetir

Aceptar Cancelar

→ *Estilo de las cabeceras*, pueden ser configuradas las fuentes, celdas y bordes.

**Editar Formato**

**Fuente**

Familia: Valor por defecto (Sistema) ▼ Tamaño: \_\_\_\_\_

Color:  Estilo: Negrita ▼ Efectos: Valor por defecto (Sistema) ▼

**Celda**

Alineación Horizontal: Centro ▼ Color de Fondo:

Alineación Vertical: Centro ▼ ☒ Ajustar Texto

**Borde**

Posición: Todo ▼ Estilo de borde: Simple ▼

Color del Borde:

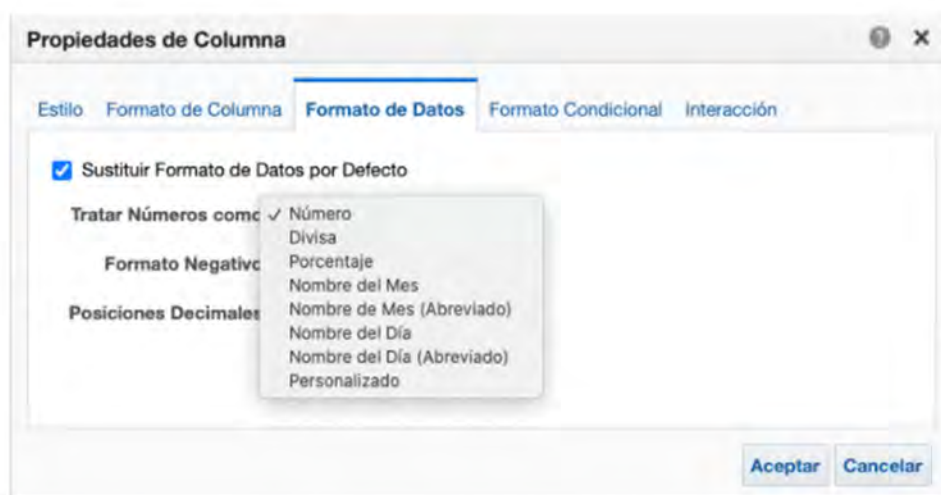
**Opciones de formato adicionales**

**Opciones de Estilo CSS Personalizado (Sólo HTML)**

Aceptar Cancelar

→ **Formato de los datos.**

- Números:** pueden configuradas las posiciones decimales, los números negativos y datos entre paréntesis.
- Divisas:** se pueden incluir los símbolos de las monedas.
- Porcentaje:** pueden configuradas las posiciones decimales, los números negativos y datos entre paréntesis.
- Nombre del mes:** es particularmente útil cuando se selecciona el criterio 'Month key', representado por números



- Personalizado:** cualquier dato puede ser configurado. Las fechas presentan diferentes formatos que pueden ser modificados utilizando la opción 'personalizado'.

→ **Formato condicional**

El formato condicional facilita la configuración de estilo para determinados valores un criterio. Formato condicional del valor 'None' del campo 'Provenance code'



Resultado del formato condicional, el valor 'None' figura en el informe destacado [azul].

Title	Encabezamientos de materia	Description	Loan Month Key	Loan Date	Hora de préstamo	Loan Circulation Desk	Provenance Code
Fundamentos de la teoría electromagnética	Electromagnetismo		January	10/01/2020	10:30	DEFAULT_CIRC_DESK	None
Campos electromagnéticos	Campos electromagnéticos; Física general		January	10/01/2020	14:27	DEFAULT_CIRC_DESK	None

→ Interacción

--Se añade una acción con el símbolo más +

Valor

Interacción Primaria Enlaces de Acción ▼

Enlaces de Acción + / X

Texto de Enlace	Acción	Mostrar Enlace
Agregar Enlaces de Acciones de Menú		

☐ No mostrar en una ventana emergente si sólo hay un enlace de acción disponible en tiempo de ejecución

☐ Activar en Totales

Especificar Canal

- Cambiar de Nivel
- Enlaces de Acción
- Enviar Eventos de Maestro-Detalle

Aceptar Cancelar

--Se añade una acción con el símbolo más +

Valor

Interacción Primaria Enlaces de Acción ▼

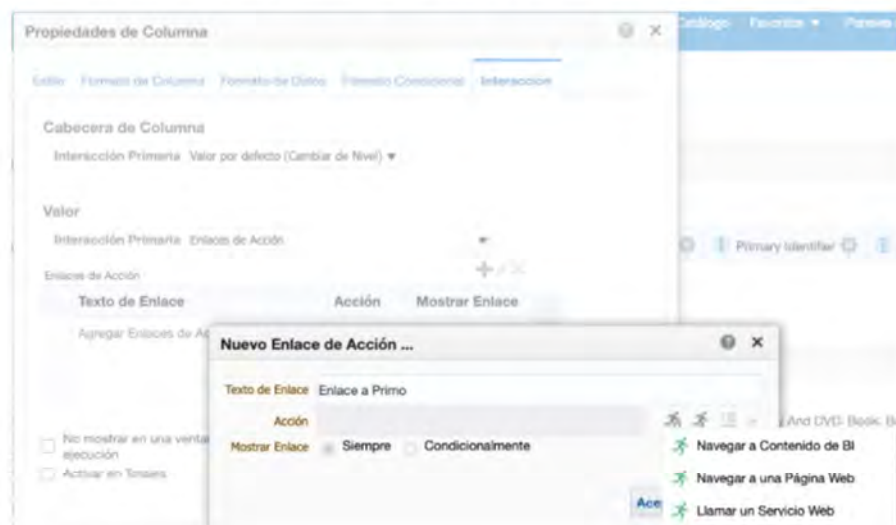
Enlaces de Acción + / X

Texto de Enlace	Acción	Mostrar Enlace
Agregar Enlaces de Acciones de Menú		

☐ No mostrar en una ventana emergente si sólo hay un enlace de acción disponible en tiempo de ejecución

☐ Activar en Totales

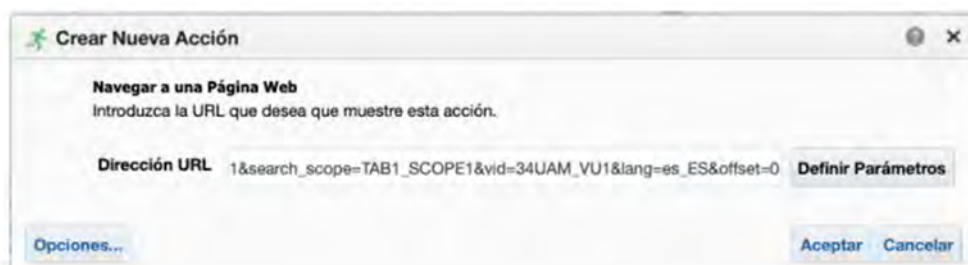
--Se añade un título al enlace y se selecciona 'Navegar a una Página Web'.



Como ejemplo, se detallan los parámetros de una biblioteca para la acción 'Navegar a una página web':

[http://bun.uam.es/primo-explore/search?query=any,contains,"MMS](http://bun.uam.es/primo-explore/search?query=any,contains,)  
Id"&tab=tab1&search\_scope=TAB1\_SCOPE1&vid=34UAM\_VU1&lang=es\_ES&offset=0

--Pulse en definir parámetros





--Configure los parámetros

**Crear Nueva Acción**

Navegar a una Página Web  
Introduzca la URL que desea que muestre esta acción.

Dirección URL  [Definir Parámetros](#)

**Definir Parámetros**

Para embeber un valor de parámetro en cualquier lugar de la URL, use el token @{Nombre} donde "Nombre" es el valor especificado en la columna Nombre de abajo.  
No hay límite para el número de parámetros que se pueden embeber.

Nombre	Petición de Datos	Valor	Fijo	Oculto	Opcional
1	query	<input type="text"/> Bibliographic Details: "MMS Id"	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
2	tab	<input type="text"/> tab1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
3	search_scope	<input type="text"/> TAB1_SCOPE1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
4	vid	<input type="text"/> S4UAM_VU1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
5	lang	<input type="text"/> es_ES	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

--La dirección URL deberá modificarla de acuerdo a esta sintaxis:

<http://bun.uam.es/primo-explore/search?>

```
query=any,contains,@{1}&tab=@{2}&search_scope=@{3}&vid=@{4}&lang=@{5}&offset=@{6}
```

Finalmente, desde el informe de Analytics al pulsar en el identificador MMS Id se podrá conectar con el registro bibliográfico en Primo.

MMS Id  
991000845839704211  
 Enlace a Primo



--Edite las propiedades del gráfico



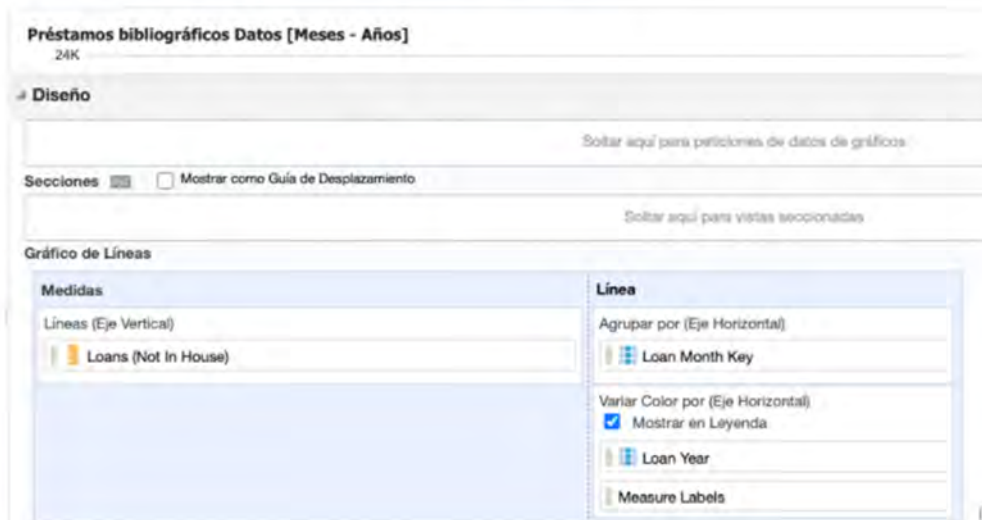
#General: sección destinada a la configuración del tamaño del gráfico, la posición que ocupa la leyenda o la inclusión de valores nulos.

#Estilo: sección del formato de los datos del gráfico, estilo 3D, y estilo del fondo y bordes.

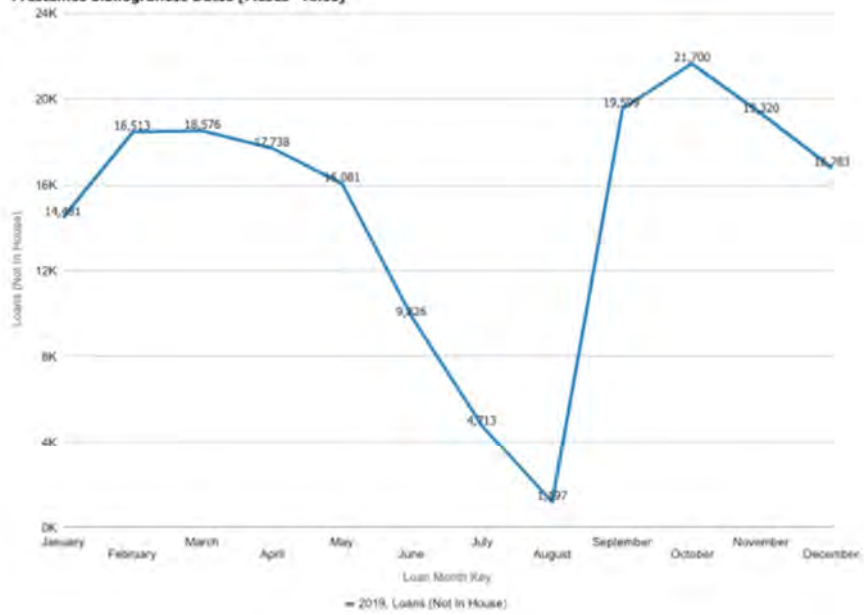
#Escala: límites y escalas de los ejes.

#Títulos y etiquetas: asignación de un título al gráfico y aplicación del diseño a marcadores de datos, leyenda y etiquetas del eje horizontal y vertical.

--Seleccione las variables. Informe sobre préstamos bibliográficos



Préstamos bibliográficos Datos [Meses - Años]



## Representaciones gráficas

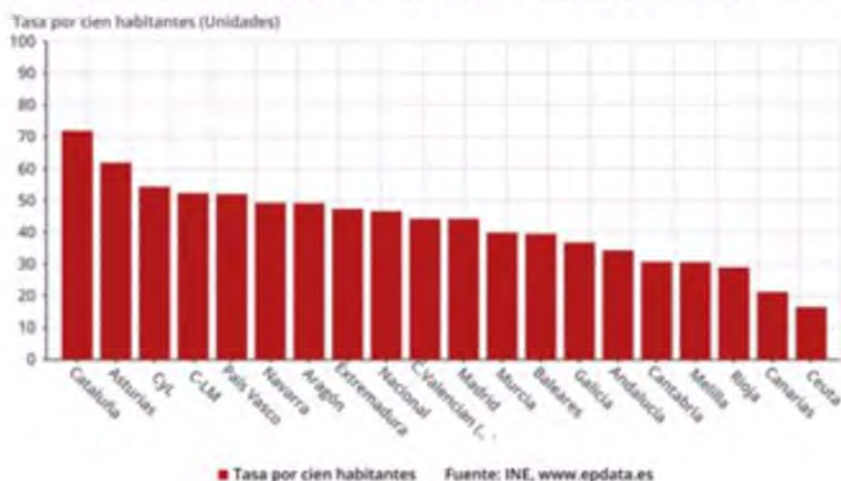
Un gráfico consiste en la representación de una distribución de datos mediante líneas, formas geométricas o puntos. Cualquier construcción gráfica se origina a partir de una tabla de datos.

Existen variedad de representaciones gráficas, pero su finalidad es que los gráficos aporten *información útil*, que ayuden a examinar los datos para describir sus características principales. Por tanto, la elección del tipo de gráfico será relevante para obtener representaciones eficaces que destaquen la información sustancial, la naturaleza de los datos y los objetivos de la investigación. Las categorías de los gráficos están en relación con los [tipos de variables](#).

### → Diagrama de barras

El diagrama de barras se utiliza para [variables cualitativas y cuantitativas de tipo discreto](#).

#### Usuarios de bibliotecas en cada comunidad autónoma

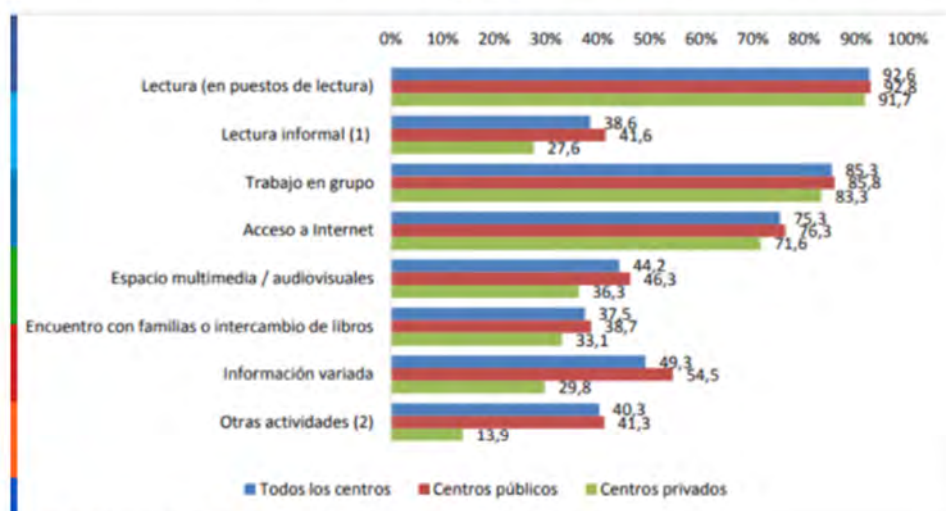


Fuente: Instituto Nacional de Estadística (INE). <https://www.ine.es>

El diagrama compara de forma rápida la frecuencia de usuarios por comunidades autónomas. Las barras del diagrama son utilizadas para comparar las distintas categorías de una variable. En general, se representan los valores de la variable separadas entre sí en el eje de abscisas y la frecuencia en el eje de ordenadas. De esta forma, para cada valor de la variable existe una barra cuya altura equivale a su frecuencia.

Variantes del diagrama de barras son los **gráficos de barras apiladas o compuestas**. Estos gráficos muestran el *tamaño relativo* (como recuento, porcentaje u otra variable numérica) de una variable de categorías, subdividida por colores en función de un subgrupo.

**Gráfico 1: Porcentaje de centros que cuentan con zonas diferenciadas en la biblioteca, por tipo de zona**

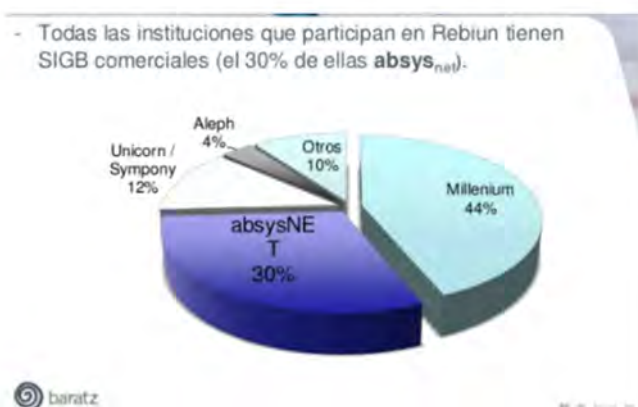


1) Lectura informal en sillones, sofás, pufs, taburetes, cojines, esterillas, etc.  
2) Seminarios y conferencias, representaciones artísticas, cuentacuentos, etc.

Fuente: Baratz: <https://www.comunidadbaratz.com/blog/el-925-de-los-centros-educativos-de-espana-dispone-de-biblioteca-escolar/>

#### → Diagrama de sectores

El diagrama de sectores se emplea cuando las variables presentan pocas categorías. El arco de cada sector equivale a la frecuencia de cada categoría. A pesar de su impacto visual su utilización es limitada, se recomienda cuando una o dos categorías dominantes abarcan gran parte del círculo y para frecuencias relativas o porcentajes.



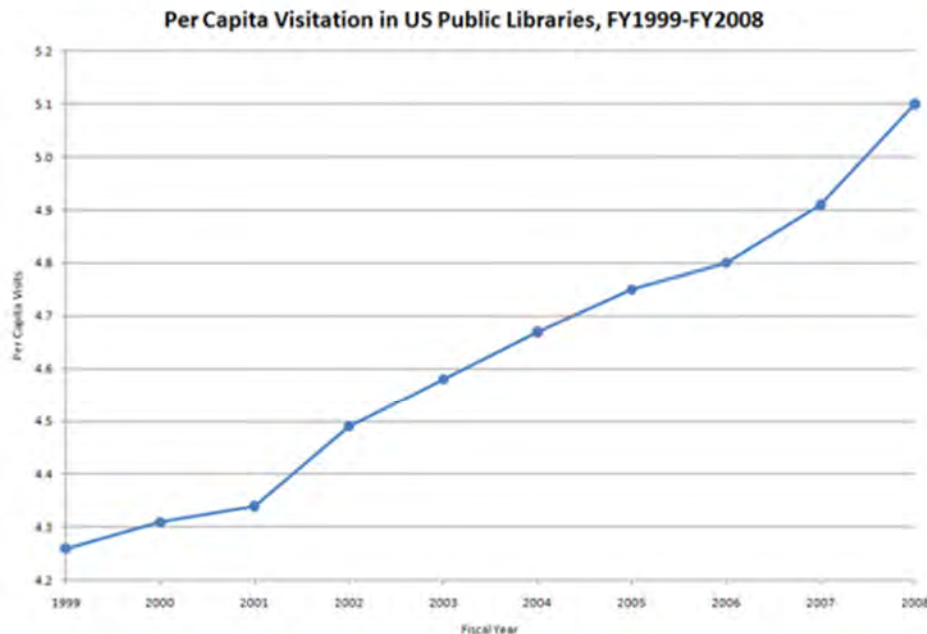
Fuente: Baratz: <https://es.slideshare.net/Baratz/soluciones-para-la-gestin-de-bibliotecas-y-redes-de-bibliotecas>

### → Series temporales

El gráfico temporal de una variable representa cada observación en relación con el momento en que se midió. La aplicación que se utiliza para las series temporales son los **diagramas de líneas**

La evolución de una cantidad a lo largo de tiempo se representa con frecuencias absolutas, porcentajes e índices.

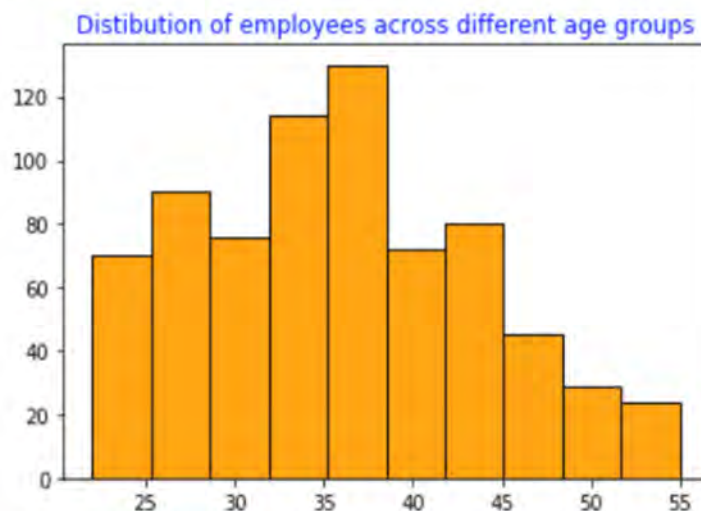
Los datos temporales del diagrama de líneas se sitúan en el eje de abscisas y las frecuencias (absolutas o relativas) en el eje de ordenadas.



Fuente: Institute of Museum and Library Services : <https://www.ims.gov/news/public-library-visits-circulation-spike-while-staff-numbers-stay-same-ims-fy2008-public>

### → Histogramas

El histograma se utiliza para **variables cuantitativas de tipo continuo**, cuando las variables contienen muchos valores y se agrupan por intervalos de los valores próximos. A diferencia del diagrama de barras, que se usa para variables discretas y por tanto, la frecuencia de cada categoría viene indicada por la altura de la barra correspondiente, el histograma representa **superficies**. Es decir, el peso de cada categoría se representa a través del área contenida en cada uno de los rectángulos. De manera que el área es proporcional a la frecuencia que representa.



Fuente: Analytics Vidhya: <https://medium.com/analytics-vidhya/python-libraries-for-data-science-day-5-matplotlib-histogramsgraph-7953cb2f8855>

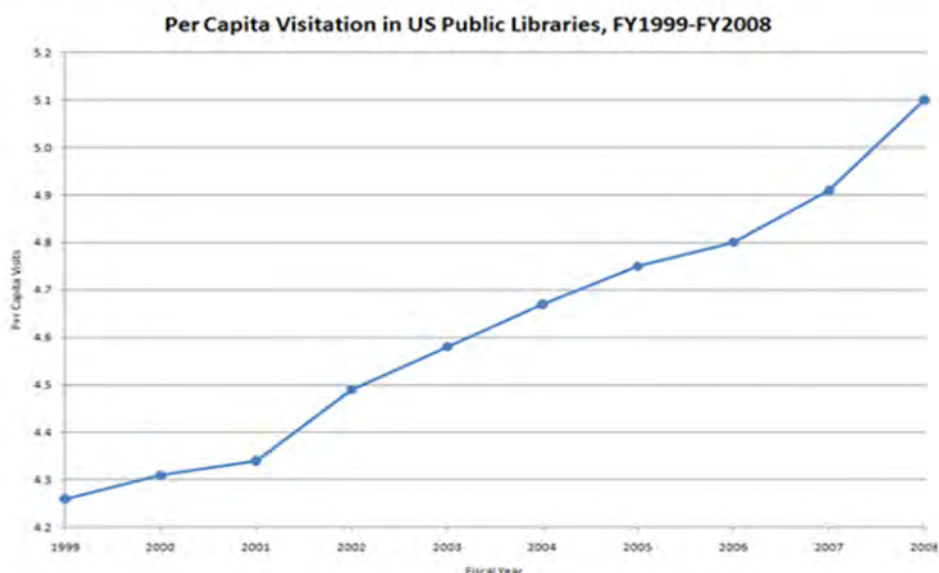


### → Series temporales

El gráfico temporal de una variable representa cada observación en relación con el momento en que se midió. La aplicación que se utiliza para las series temporales son los **diagramas de líneas**

La evolución de una cantidad a lo largo de tiempo se representa con frecuencias absolutas, porcentajes e índices.

Los datos temporales del diagrama de líneas se sitúan en el eje de abscisas y las frecuencias (absolutas o relativas) en el eje de ordenadas.



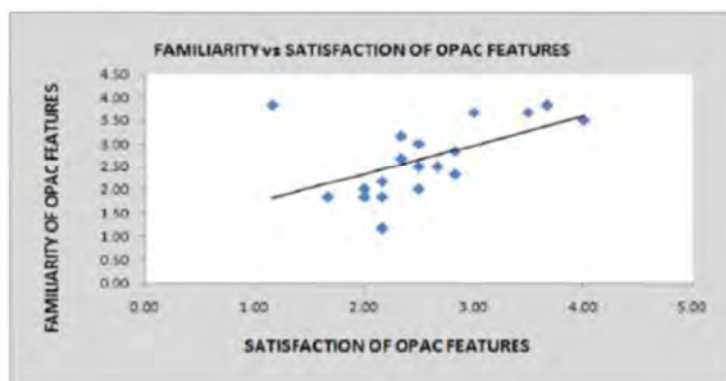
Fuente: Institute of Museum and Library Services : <https://www.ims.gov/news/public-library-visits-circulation-spike-while-staff-numbers-stay-same-ims-fy2008-public>

### → Nube de puntos

Cuando lo que interesa es representar la distribución conjunta de dos variables para observar la relación que existe entre ellas se utiliza el **diagrama de dispersión o nube de puntos**.

Este tipo de gráficos muestra los valores de la variable mediante un eje cartesiano de dos perpendiculares. El eje X, abscisas, representa la variable independiente o 'causativa', y el eje Y (ordenada) representa la variable dependiente o 'efecto'.

Aparte de representar los patrones conjuntos de dos variables, los diagramas de nube de puntos son útiles para detectar *outliers*, o valores que se alejan extremadamente de la tendencia central de los datos de la distribución.



Fuente: Zabala, J.L., Deonoso, T., Bayarcal, A. (2017). Antecedents of Searching Satisfaction: A Usability Study of an Online Public Access Catalog. *ASLP Journal* (3).

(1) Camarero, L.A. (2010). *Estadística para la investigación social* (UNED). Málaga: Garceta,

ExLibris: [creación de nuevos informes](#),



### 3. Filtros

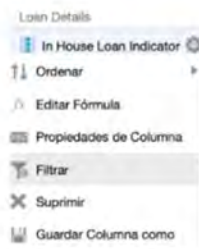
- [Introducción](#)
- [Funciones SQL](#)
- [Material complementario](#)

#### Introducción

Los filtros permiten limitar la cantidad de datos desplegados en el informe y se aplican antes de su ejecución. Los filtros afectan al contenido del informe, y por tanto a los resultados.

Pasos a seguir:

- a. Haga click en el icono de 'Opciones de la columna' cuyos datos se quieren filtrar.



- b. Seleccione la opción(es) de la columna desde el listado 'Operador'



-El campo '**se solicita**' selecciona todos los valores del criterio y es necesario en los diseños del 'Panel de control'.

Una fórmula más avanzada y precisa para filtrar la información consiste en la utilización de las funciones SQL.

## Funciones SQL

El lenguaje de consulta estructurado SQL (Structured Query Language) permite especificar diferentes operaciones en bases de datos relacionales. El lenguaje SQL está compuesto por comandos y funciones de agregado. Estos elementos se combinan para realizar operaciones automáticas en la base de datos.

Esta sección contiene las fórmulas SQL frecuentemente utilizadas en Alma-Analytics y Visualización de datos.

La tabla de funciones SQL que se expone a continuación, está estructurada en 4 secciones:

- Funciones por categorías
- Sintaxis SQL
- Descripción de las funciones
- Filtro de un criterio y funciones (fx), fórmula de un criterio.

### 1. Filtro de un criterio



### 2. Funciones (fx), fórmula de un criterio.

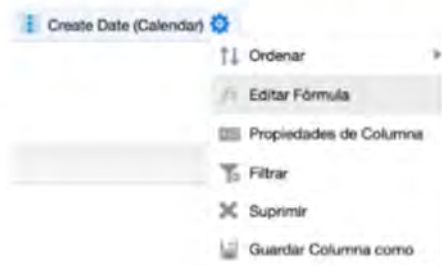


Tabla de instrucciones SQL

Funciones temporales	Sintaxis	Descripción	Filtro (1) Funciones (2)
<b>CURRENT_DATE</b>	<b>CURRENT_DATE</b>	<b>Función que actualiza los datos a la fecha actual</b>	1
	"Loan Date"."Loan Month Key" = MONTH(CURRENT_DATE)-1	Préstamos, mes anterior a la fecha actual.	1
	"Loan Date"."Year Key" = YEAR(CURRENT_DATE)	Préstamos, año actual.	1
	YEAR("Bibliographic Details"."Creation Date") = YEAR(CURRENT_DATE)-1	Creación de los registros bibliográficos, un año anterior al actual.	1
	YEAR("Invoice Line"."Invoice-Creation Date") = YEAR(CURRENT_DATE) AND MONTH("Invoice Line"."Invoice-Creation Date") = MONTH(CURRENT_DATE)	Líneas de facturas creadas, el mes actual del año actual.	1
<b>TIMESTAMPADD (SQL_TSI_Intervalo)</b>	<b>TIMESTAMPADD (SQL_TSI_Intervalo)</b> Intervalos: DAY, WEEK, MONTH, YEAR, HOUR	<b>Función que modifica la expresión fecha/hora de acuerdo al número especificado.</b>	1
	"Loan Date"."Loan Date" >= TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Préstamos, 7 días antes de la fecha actual.	1
	"Loan Date"."Loan Date" <= TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Préstamos, antes de 7 días respecto a la fecha actual.	1
	"Due Date"."Due Date" = TIMESTAMPADD(SQL_TSI_DAY, +2, CURRENT_DATE)	Devoluciones debidas, dos días en adelante a la fecha actual.	1
	"Loan Date"."Loan Date" BETWEEN TIMESTAMPADD(SQL_TSI_DAY, -30, CURRENT_DATE) AND (CURRENT_DATE)	Préstamos producidos, entre la fecha actual y 30 días antes.	1
<b>TIMESTAMPDIFF(interval , expr, timestamp2)</b>	Intervalos: DAY, WEEK, MONTH, YEAR Expr: primer criterio Timestamp2: segundo criterio (date), o CURRENT_DATE.	<b>Función que calcula la diferencia de fechas en: años, meses, semanas, días u horas.</b> <b>Analytics:</b> -Editar fórmula (fx) -Calendario/Fecha -TimestampDiff	2
	TIMESTAMPDIFF(SQL_TSI_DAY, "Physical Item Details"."Creation Date", "Physical Item Details"."Modification Date")	Valor calculado. Diferencia en días desde la fecha de creación de ítem y la fecha de su modificación.	2
	TIMESTAMPDIFF(SQL_TSI_DAY, "Physical Item Details"."Last Loan Date", CURRENT_DATE)	Valor calculado. Diferencia en días desde la última fecha de préstamo y la fecha actual.	2

Funciones de agregación	Sintaxis	Descripción	Filtro (1) Función (2)
<b>Avg</b>	Avg(TIMESTAMPDIFF(SQL_TSI_DAY,"PO Line"."Sent Date","PO Line"."Receiving Date (Latest in POL)"))	Media de días entre: -Fecha del envío POL -Fecha de recepción POL <b>Analytics:</b> -Editar fórmula (fx) -Agregado -Avg	2
<b>CountDistinct</b>	COUNT(DISTINCT "Bibliographic Details"."MMS Id")	Transforma identificadores distintivos como: MMS-ID, Holding-ID, Borrower details. Primary identifier, en números enteros. Con esta función datos categóricos se convierten en cuantitativos y pueden ser tratados como números. Por tanto, se pueden realizar operaciones aritméticas. <b>Analytics:</b> -Editar fórmula (fx) -Agregado -CountDistinct	2
<b>SUM</b>	SUM("Fund Transactions"."Transaction Amount" by "Vendor"."Vendor Name")	Calcula la suma por valores <b>Analytics:</b> -Editar fórmula (fx) -Agregado -Sum	2
<b>TopN</b>	TOPN("Loan"."Loans (Not In House)", 10)	Los 10 títulos más prestados el último mes. <b>Analytics:</b> -Editar fórmula (fx) -Agregado -TopN	2
<b>BOTTOMN</b>	BOTTOMN("Loan"."Loans (Not In House)", 10)	Recupera los títulos menos prestados el último mes. <b>Analytics:</b> -Editar fórmula (fx) -Agregado -Bottomn	2



Cadena	Sintaxis	Descripción	Filtro (1) Función (2)
<b>SUBSTRING</b>	SUBSTRING("Loan Date"."Loan Time" FROM 0 FOR 5)	Permite reducir el número de caracteres de un criterio. Devuelve las horas con la expresión de 5 caracteres. Por ejemplo: 09:15 <b>Analytics:</b> -Editar fórmula (fx) -Cadena -Substring	2
Expresiones	Sintaxis	Descripción	Filtro (1) Función (2)
<b>FILTER</b>	<b>FILTER</b> ("Physical Item Details"."Num of Items" <b>USING</b> ("Physical Item Details"."Process Type" = 'Acquisition'))	Extrae un valor numérico. Requiere un primer valor de medida y un segundo criterio categórico sobre el que obtendrá un dato numérico. En este caso, del área Physical items se extraen del número de ítems existentes, los que están en una nueva fase de adquisición.	2
<b>CASE</b>	<b>CASE</b> "Bibliographic Details"."Material Type" <b>WHEN</b> NULL <b>THEN</b> 'SIN DATOS' <b>ELSE</b> "Bibliographic Details"."Material Type" <b>END</b>	Secuencia lógica: <b>CASE:</b> caso <b>WHEN:</b> cuando ocurre algo <b>THEN:</b> entonces 'expresión'. <b>ELSE:</b> sino ocurre entonces 'expresión'. <b>END:</b> fin. Ejemplo. Si el 'tipo de material' está vacío, entonces escribe 'Sin datos'. De lo contrario, escribe el 'tipo de material'.	2
	<b>CASE</b> "Bibliographic Details"."Bibliographic Level" <b>WHEN</b> 'a' <b>THEN</b> 'Monographic component part' <b>WHEN</b> 'b' <b>THEN</b> 'Serial component part' <b>WHEN</b> 'c' <b>THEN</b> 'Collection' <b>WHEN</b> 'd' <b>THEN</b> 'Subunit' <b>WHEN</b> 'i' <b>THEN</b> 'Integrating resource' <b>WHEN</b> 'm' <b>THEN</b> 'Monograph/Item' <b>WHEN</b> 's' <b>THEN</b> 'Serial' <b>ELSE</b> "Bibliographic Details"."Bibliographic Level" <b>END</b>	Bibliographic details. Bibliographic level. Leader, position: 07	2

Campos bibliográficos	Sintaxis	Descripción	Filtro (1) Función (2)
<b>LIKE, IS NOT LIKE</b> % -	"Bibliographic Details"."Title" LIKE 'Manual de%'	<b>LIKE</b> , devuelve los términos especificados. <b>IS NOT LIKE</b> , excluye los términos especificados. <b>%</b> , trunca varios caracteres desde la raíz. <b>-</b> , trunca un carácter.	

#### Material complementario

[Ex Libris: Presentations and Documents - Analytics. Advances functions](#)

[Rogers, M: Tips and tricks for Alma Analytics](#)



## 4. Paneles de control (Dashboards)

- Introducción
- Creación de paneles de control
- Petición de datos de panel de control
  - --Selección del área temática
  - --Salvar el panel de control
  - --Criterios del panel de control
  - --Agregar criterios al Panel de control
  - --Composición del panel de control
- Paneles de control temático
  - --Añadir más de un informe a un panel por área temática
- Panel de control principal

## Introducción

Los paneles de control se utilizan en Alma Analytics para reunir colecciones de informes de las mismas áreas temáticas, y para filtrar información de acuerdo a determinados parámetros.

Los paneles de control responden a tres categorías:

a) **Panel de control principal de la biblioteca**, también denominado '**Paneles de Control**', que reúne todos los paneles de control por áreas temáticas de la biblioteca.



b) **Paneles de control por áreas temáticas**, pueden formar parte del panel del control principal de la biblioteca, del área privada 'Mis carpetas', o de otras carpetas del directorio.

La siguiente imagen muestra cuatro paneles de control por áreas temáticas (Adquisiciones, Analytics Panel de Control, Circulación, Estadísticas de circulación), dependientes del 'Panel de control principal' o 'Paneles de control'.



c) **Petición de datos de Panel de control**. Se trata de paneles de control asociados a informes individuales y sirven para filtrar la información por determinados parámetros.

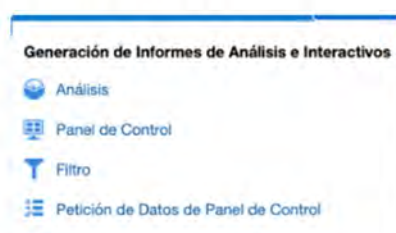


## Creación de paneles de control

Un panel de control siempre está unido a un informe o grupo de informes. Para que funcione un panel de control su área temática deberá ser la misma que la del informe o colección de informes.

Desde el menú superior de Alma Analytics se pulsa en **Crear**, se despliega un listado que incluye dos tipos de paneles de control:

- 'Panel de control', agrupa temáticamente a varios informes.
- 'Petición de Datos de Panel de Control', asociado a un informe



## Petición de datos de panel de control

La primera información que se muestra al pulsar en 'Petición de Datos de Panel de Control' es la selección del área temática.

--Selección del área temática

**i** El área temática de la 'Petición de datos de panel de control' y la del informe asociado debe ser la misma.

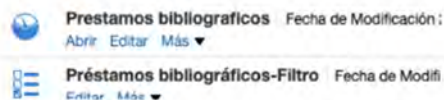


--Salvar el panel de control

Se recomienda, para minimizar la pérdida de datos, guardar el panel de control con un nombre. Los iconos de 'salvar' y 'salvar como' se encuentran en la parte superior de la página.



Se sugiere guardar el panel con el mismo nombre del informe más un sufijo, por ejemplo panel o filtro, y archivarlo en la misma carpeta.



--Criterios del panel de control

A continuación, se agregarán al panel de control los criterios del informe de los que se quiere filtrar información.

Etiqueta de Petición de Datos	Tipo	Petición para
Página 1	Página	
Library Name	Valor de columna	Library Name
Location Name	Valor de columna	Location Name
User Group	Valor de columna	User Group
Title	Valor de columna	Title

**i** Los **criterios del informe** que vayan a formar parte del 'Panel de control' deberán incluir **el operador del filtro 'Se solicita'**, que selecciona automáticamente todos los valores del criterio, o bien **una selección de variables del criterio**.

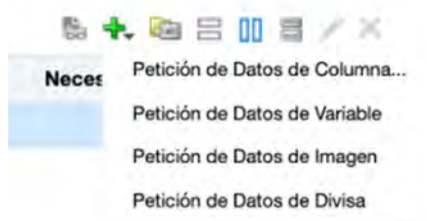


## Filtros del informe.



## --Agregar criterios al Panel de control

Con el ícono 'más' se añade una petición de datos de columna o criterio.



Ejemplo de configuración de un criterio en el Panel de control:

*Criterio:* 'Library name'.

*Etiqueta:* el nombre del criterio puede ser modificado pulsando de 'Etiqueta personalizada'.

*Descripción:* se puede añadir una breve explicación al criterio.

*Operador:* contiene las mismas opciones de selección que las del criterio del informe.

*Valores de la lista de opciones:* se pueden seleccionar las variables que interesen.

**Nueva Petición de Datos:**

**Petición para Columna** Item Location at time  $\leq$  3.

**Etiqueta** Library Name

☐ Etiqueta Personalizada

**Descripción**

**Operador** es igual que / está en

**Entrada de Usuario** Lista de Opciones

**Opciones**

**General Más**

**Valores de Lista de Opciones**

- Ninguna
- Todos los Valores de Columna
- ☒ Valores de Columna Específicos
- Resultados de SQL
- Miembros de Grupos
- Todos Valores Columna/Grupos Específicos

☐ Incluir la opción "Todos los Valores de Columna" en la lista

☒ Permitir al usuario seleccionar varios valores

☒ Permitir al usuario escribir valores

☐ Requerir entrada de usuario

☐ Limitar valores por

**Rellenar Automáticamente** Ninguna  
Valor Disponible en Primer Lugar

**Selección por defecto** Ninguno

**Definir una variable** Ninguna

**Aceptar Cancelar**

#### --Composición del panel de control

Los criterios se pueden reordenar con las flechas que figuran en el margen derecho de la pantalla. Además, cuando se marcan los elementos de acción de la 'nueva columna', la ordenación de los criterios cambia de posición en el área 'Mostrar'.

Etiqueta de Petición de Datos	Tipo	Petición para	Descripción	Necesario	Nueva Columna
Página 1	Página				
Library Name	Valor de columna	Library Name			<input type="checkbox"/>
Location Name	Valor de columna	Location Name			<input type="checkbox"/>
User Group	Valor de columna	User Group			<input checked="" type="checkbox"/>
Title	Valor de columna	Title			<input checked="" type="checkbox"/>

**Mostrar**

Página 1

Library Name --Seleccionar Valor-- User Group --Seleccionar Valor-- Title contiene cualquiera --Seleccionar Valor-- Location Name --Seleccionar Valor--

Aplicar Restablecer

### Paneles de control temático

Como se ha indicado anteriormente, se creará un 'Panel de control temático' dependiente del 'Panel de control principal' para poder diseñar posteriormente un 'Portal de Analytics'. El permiso para realizar esta tarea corresponde al 'Administrador de Analytics'.

#### --Paneles de Control

Sitúe el cursor en '**Paneles de Control**'



#### --Panel de control temático

Desde el menú '**Crear**' de Alma-Analytics seleccione '**Panel de control**'. Introduzca un nombre identificativo del área temática del panel de control.

**Nuevo Panel de Control**

Seleccionar un nombre y una ubicación para el nuevo panel de control

Nombre

Descripción

Ubicación

Contenido ☒ Agregar contenido ahora ☐ Agregar contenido más tarde (Crear panel de control vacío)

Aceptar Cancelar

A continuación, se abrirá la página de configuración del panel de control.

Pulse en '**Catálogo**' y compruebe la ubicación del panel de control 'Préstamos bibliográficos' en el directorio de '**Paneles de control**'





Identifique la '**página 1**' con un nombre [Préstamos bibliográficos], desde **Mas**.



Edite la página 'Préstamos bibliográficos'

Desde '**catálogo**' seleccione en el directorio la carpeta que contiene el **informe** '**Préstamos bibliográficos**' y la petición de datos de control '**Préstamos bibliográficos-Filtro**' .

Arrastre al lienzo en primer lugar '**Préstamos bibliográficos-Filtro**', y a continuación '**Préstamos bibliográficos**'.



Guarde el panel y ejecute la acción. Automáticamente se desplegará el panel de control.

Library Name: Item location --Seleccionar Valor--

Location Name --Seleccionar Valor--

Loan Year 2020

Loan Month Key --Seleccionar Valor--

Loan Date Entre

Title --Seleccionar V--

Encabezamientos de materia --Seleccionar V--

Material Type: Item --Seleccionar Valor--

Provenance Code --Seleccionar V--

Circ Desk Name --Seleccionar Valor--

User Group --Seleccionar Valor--

Facultad - Usuario --Seleccionar Valor--

Usuarios - Fechas de expiración Entre

Loans (Not in House) 68,680

Núm. de usuarios 12,286

Library Name, Item location	Location Name	Permanent Call Number	Barcode	Title	Encabezamientos de materia	Description	Loan Month Key	Loan Date	Hora de préstamo	Loan Circulation Desk	Provenance Code	Material Type Item	Material Type Bibliographic	Item Policy
Cartoteca	Cartoteca	CART/MTN50-PLEG	5405745168	Comillas	Sin datos		January	10/01/2020	12:02	DEFAULT_CIRC_DESK	Compra	Map	Map	Semi
	Cartoteca	CART/MTN50-PLEG	5409542082	Foz	Sin datos		January	10/01/2020	11:26	DEFAULT_CIRC_DESK	Compra	Map	Map	Semi
	Libre acceso	CART/551.4/MUÑ	5411741637	Geomorfología general	Geomorfología; Geología general; Geografía		January	17/01/2020	10:31	DEFAULT_CIRC_DESK	Compra	Book	Book	Semi
	Cartoteca	CART/ArmaríoGeológicos3	5406554194	Almagro	Sin datos		January	20/01/2020	10:03	DEFAULT_CIRC_DESK	Compra	Map	Map	Semi

--Añadir más de un informe a un panel por área temática

a) Edite el panel temático [Préstamos bibliográficos]

--Añadir más de un informe a un panel por área temática

a) Edite el panel temático [Préstamos bibliográficos]

**Carpetas**

- MATERIALES ESPE
- PRESTAMOS
- Préstamos bibliográficos**
- REFRIN

**Tipo Todo**

**Ordenar Nombre**

dashboard layout Fecha de Modificación 22/11/2

Propiedades Más

**Préstamos bibliográficos** Fecha de Modificació

Abrir Editar Más

b) Pulse en el primer icono 'Agregar pagina de panel de control'



Vista Previa

Ejecutar

c) Asigne un título a la nueva página del panel de control.

**Agregar Página de Panel de Control**

Nombre de Página Autopréstamos

Descripción de Página

Aceptar Cancelar

d) Pulsando en 'Aceptar', se abrirá el panel de control para la configuración de la nueva página [Autopréstamos]. Desde 'catálogo' se seleccionarán los informes con los filtros correspondientes.

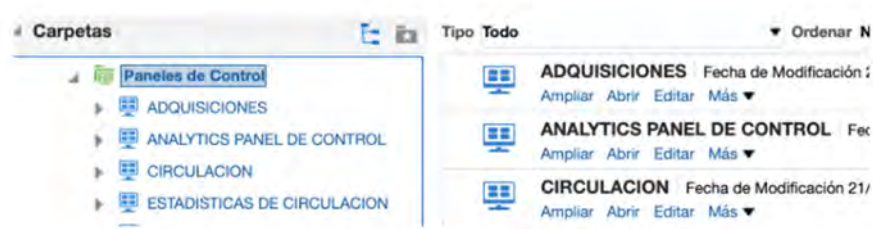


### Panel de control principal

El panel de control principal reúne el conjunto de los paneles por áreas temáticas, así como a los informes integrados en los paneles temáticos.

**i El panel de control principal da acceso al conjunto de los informes que serán mostrados al personal de la biblioteca.**

El panel de control principal, deberá incluir un panel temático destinado a la creación del '**Portal de Analytics**'. En la imagen inferior, el panel '**Analytics Panel del Control**' está configurado como un portal.



## 5. Portal de Analytics

- Configuración del Portal de Analytics
- Portal de Analytics de la Biblioteca UAM
- Informes descriptivos por áreas temáticas
  - → Adquisiciones:
  - → Circulación:
  - → Mantenimiento de colecciones
  - → Registros bibliográficos, ítems y portfolios creados
- Estadísticas por áreas temáticas
  - → Adquisiciones
  - → Recursos electrónicos
- Rebiun (Estadísticas anuales)

El portal de Analytics es una aplicación que ofrece acceso a una información centralizada e integrada de las actividades que se llevan a cabo en las bibliotecas a través de una interfaz única. El portal presenta gran cantidad de información actualizada, organizada, detallada y fácil de usar.

### Configuración del Portal de Analytics

Desde el panel de control principal de la Universidad cree un panel de control temático con un título distintivo.

a) Arrastre al lienzo las columnas y secciones



b) A continuación, agregue un 'enlace a la imagen' para añadir un logo al panel.

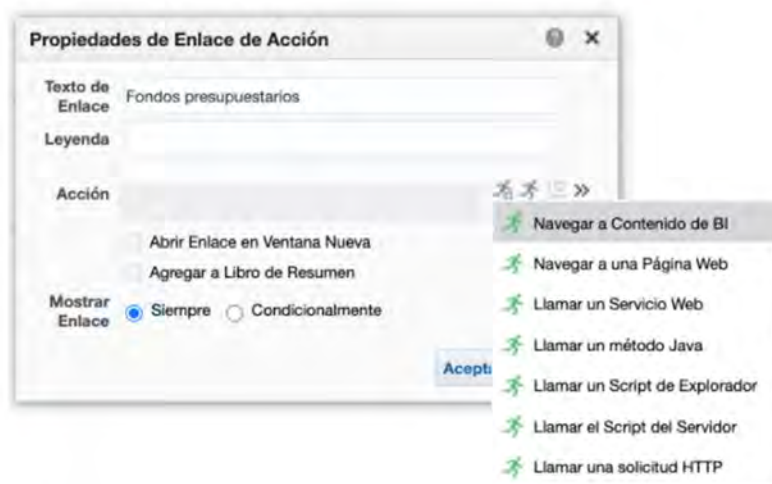




c) Pulse en las propiedades de 'Texto' para introducir la narrativa de la página:

d) Utilice columnas para crear secciones y espacios intermedios.

- e) Para configurar una sección [Adquisiciones] introduzca el título con el objeto **'Texto'**.  
Para enlazar añadir un apartado [Fondos presupuestarios] utilice el objeto **'Enlace de acción'**



--Enlace con Alma. Cuando haya completado el diseño de la página web, deberá enlazarla con Alma a través del objeto **'Dashboard'**, incluido en **'Analytics Objects List'**.

Analytics  
Design Analytics  
Data Visualization  
Subscribe to Analytics

Analytics Objects  
Analytics Objects List  
Analytics User Statistics  
Analytics Citation Attribute Types

Reports  
ANALYTICS. PANEL DE CONTROL

ANALYTICS. PANEL DE CONTROL

--Configure el objeto **'Dashboard'**. Asigne diferentes roles para que el portal pueda ser consultado por el personal de la biblioteca.

Alma Analytics Object Cancel Save

Portal de Analytics Info

General Information

Title \* Portal de Analytics

Analytic Folder \* Universidad Autónoma de Madrid/Paneles de Control Name \* Portal de Analytics

Type \* Dashboard

Description

Roles Add Role

Role	Description
1 Analytics Administrator	Manages Alma Analytics configuration including Analytics Objects List, Analytics User Statistics, and Analytics Citation Attribute Types
2 Circulation Desk Operator	Manages circulation desk actions such as loan, return, renew etc

## Portal de Analytics de la Biblioteca UAM

El modelo de portal que se detalla a continuación está estructurado a partir de tres páginas web:

1. Informes descriptivos
2. Estadísticas por áreas temáticas
3. Rebiun (estadísticas anuales)



## Informes descriptivos por áreas temáticas

La página web de 'Informes descriptivos' está organizada en cuatro áreas temáticas:

### → Adquisiciones:

Contiene informes con datos procedentes del área 'Funds expenditure'.

Cada informe incluye filtros que permiten realizar búsquedas avanzadas por varios criterios simultáneamente.

Para el conocer la configuración del informe y los filtros de búsqueda avanzada véase [paneles de control temáticos](#)

Fondos presupuestarios

Fondos presupuestarios

Fiscal Period Description: FY 2020

Vendor Name: [Seleccionar Valor]

Library Name: [Seleccionar Valor]

Order Line (Material Type): [Seleccionar Valor]

Fund Ledger Name: [Seleccionar Valor]

Material Type (POL): [Seleccionar Valor]

Invoice Status (POL): [Seleccionar Valor]

Material Type (Bibliographic): [Seleccionar Valor]

Acquisition Method Description: [Seleccionar Valor]

Site: [Seleccionar Valor]

Aplicar Restricciones

Datos filtrados por el año fiscal:

Total: 1,217,198.296

Library Name	Fund Ledger Name	Vendor Name	Title	Total	Acq.	Ext.	Quantity for pricing	Invoice-REF Type	Invoice Number	Invoice Date	Order Line Type	Reporting Code Description	Material Type (Bibliographic)	Material Type (POL)	PO Number	PO Line Reference	Invoice Status	Acquisition Method Description	Notes (Bibliographic)	Notes
Chemistry	Biblioteca Central	CAMBRIDGE UP	Gravity Resonance and Quantum Entanglement	175,800	175,800	8,800	1	Line Extension	121000000	28/08/2019	Electronic Book (On-line)	Electronic Book	Book	Book	POL-624135	POL-21044	Paid Invoiced	Purchase of Vendor System	12101121000000000000	12101121000000000000

→ **Circulación:**

Contiene informes del área temática 'Fulfillment'. Incluye informes de ítems en cuarentena, devoluciones, préstamos, reclamaciones, reservas, top préstamos y reservas y tránsitos entre bibliotecas.

Como en el apartado anterior, cualquier informe contiene un **panel de control temático** que facilita las búsquedas avanzadas.

Préstamos														
Préstamos bibliográficos    Préstamos No bibliográficos    Préstamos por autoconsumo    Préstamos por movilidad														
Préstamos bibliográficos														
Library Name, Item Location: [Seleccionar Valor] Location Name: [Seleccionar Valor] Loan Year: 2020 Loan Month Key: [Seleccionar Valor] Loan Date Range: [Seleccionar Valor] Encabezamientos de materia: [Seleccionar Valor] Material Type, Item: [Seleccionar Valor] Provenance Code: [Seleccionar Valor] Circ Desk Name: [Seleccionar Valor] User Group: [Seleccionar Valor] Facultad - Usuario: [Seleccionar Valor] Usuario - Fecha de expiración Ende: [Seleccionar Valor]														
Loans (Not In House): 69,587      Mm. de usuarios: 12,358														
Library Name, Item Location	Location Name	Permanent Call Number	Barcode	Title	Encabezamientos de materia	Description	Loan Month Key	Loan Date	Hora de préstamo	Loan Circulation Desk	Provenance Code	Material Type Item	Material Type Bibliographic	Reg. Policy

→ **Mantenimiento de colecciones**

Esta sección esta categorizada en tres apartados:

- Colección electrónica: contiene inventarios e informes específicos de libros y revistas electrónicas.
- Colección en papel: contiene inventarios, informes de revistas en papel y un informe destinado a expurgos.
- Colección mixta en papel y formato electrónico: libros y revistas.

Como el resto de los informes, incluyen el área de búsquedas avanzadas.

→ **Registros bibliográficos, ítems y portfolios creados**

Este apartado recupera los registros bibliográficos, ítems y portfolios creados por las bibliotecas en los dos últimos años.

Estos informes se distribuyen por suscripción mensualmente a través de *Analytics Objects List*.

### Estadísticas por áreas temáticas

Esta página web recopila los datos estadísticos correspondientes a los informes descriptivos.

La mayor parte de los informes contienen filtros avanzados de búsqueda y gráficas dinámicas que mejoran la visualización de los datos estadísticos.

Esta web está dividida en tres áreas temáticas:

1. Adquisiciones
2. Circulación
3. Recursos electrónicos



#### → Adquisiciones

Contiene informes sobre el 'Estado de cuentas', con información detallada de los fondos asignados, gasto realizado, gasto comprometido y asignación restante.

Otro apartado de esta sección son los informes de 'Proveedores', se incluyen informes detallados de compras y uno sobre evaluación de los proveedores basados en el intervalo entre peticiones y entregas.

Un tercer apartado está dedicado a los datos estadísticos de adquisiciones por tipos de materiales.

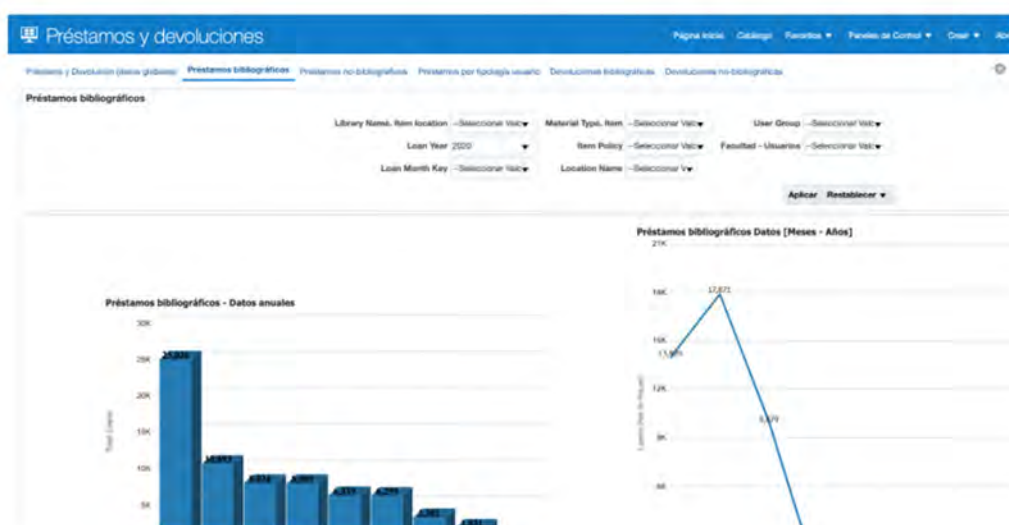
#### → Circulación

Comprende informes estadísticos de circulación con gráficas añadidas.

Este apartado se subdivide en

--Préstamos y devoluciones: la estructura del informe incluye, una tabla dinámica con los datos, búsquedas avanzadas y dos gráficas, una acumulativa por bibliotecas y otra temporal por meses.





Además, en este apartado se pueden encontrar informes sobre 'Renovaciones y auto-renovaciones', 'Reservas' y 'Tránsitos entre bibliotecas'.

#### → Recursos electrónicos

Se incluyen informes sobre 'Coste-uso de los recursos electrónicos', colecciones electrónicas y portafolios y un informe sobre los 'Link resolvers' de las colecciones electrónicas:

**Colecciones electrónicas. Link ...**

Página Inicio | Catálogo | Frecuentados | Perfiles de Control | Cerrar | Ayuda

**Colecciones electrónicas. Link Resolver**

Colección, Open URL --Seleccionar Val--  
 Title --Seleccionar Val--  
 Request Year 2019 --Seleccionar Val--

Aplicar Restablecer

Selección. Número clics >=1  
 Nóm. Títulos 15.382

Sumar	Título	Enlace al registro en BUNE	Number of Requests	Number of Clicked Requests	Number of Requests Without Services	% Requests Without Services from Requests	Request Year
Muñoz, Elena	1519 - anuario de la Sociedad Española de Literatura General y Comparada	991007629785404211	1	1	0	0.00%	2019

## Rebiun (Estadísticas anuales)

Esta página web recopila las estadísticas anuales retrospectivas de la biblioteca, de acuerdo con los indicadores de la plantilla de datos de Rebiun.

Los apartados de la plantilla Rebiun analizados son:

- 5. Colecciones
- 6. Servicios
- 9. Gasto

**ANALYTICS PANEL DE CONTROL**

INFORMES DESCRIPTIVOS [2019-2020]
ESTADÍSTICAS POR ÁREAS TEMÁTICAS [2019-2020]
VISUALIZACIÓN DE DATOS
**REBIUN (ESTADÍSTICAS ANUALES)**

**FORMULARIO ESTADISTICO REBIUN 2019**  
Se presentan los datos de las bibliotecas UAM correspondientes a los indicadores estadísticos **REBIUN [2019]**, apartados:  
**5. Colecciones**  
**6. Servicios**  
**9. Gasto**  
La información completa de las estadísticas de las bibliotecas universitarias REBIUN se puede encontrar en la dirección:  
<https://rebiun.um.es/rebiun/admin/ManageIndicatorsPage>

**APARTADO 5: COLECCIONES**

**APARTADO 6: SERVICIOS**

**APARTADO 9: GASTO**

Se presenta a continuación el informe del indicador 5.1.1. 'Títulos de monografías en papel antes de 01/01/2020', basado en el área temática 'Titles'. Contiene una sección de búsqueda y una breve descripción del indicador.

**5.1.1. Títulos de monografías e...**

PÁGINA INICIAL
Catálogo
Favoritos
Panel de Control
Crear

**5.1.1. Títulos monografías en papel antes del 01/01/2020**

Para consultar el número de títulos por biblioteca, se seleccionará desde el campo: 'Library Name (Physical)' el existente: 'Búsqueda - (Filtros)'. La opción: 'Contiene: nombre de la biblioteca'.

Library Name (Physical) --Seleccionar Valor--

Title Creation Year --Seleccionar Valor--

Aplicar
Restablecer

**REBIUN 5.1.1. "Número total de títulos (registros bibliográficos) de monografías en papel informatizados antes del 01/01/2020". Rectificable si ha habido expurgo.**  
El número de títulos por biblioteca indica los títulos únicos y comunes con otras bibliotecas. El dato real de los títulos únicos de las bibliotecas de la UAM es el total de títulos.

Library Name (Physical)	Resource Type	Material Type (Bibliographic)	Num of Titles (All)	Num of Physical Items (Active)	Title Creation Year
Cartoteca	Book - Physical	Book	1,045	1,038	2018
			3	3	2019
	Manuscripts - Physical	Book	1	1	2018
	Undefined	Book	7	7	2018
Cartoteca: Ciencias	Book - Physical	Book	49	191	2018
Cartoteca: Ciencias; Derecho; Económicas; Educación; Humanidades; Psicología	Book - Physical	Book	1	19	2018
Cartoteca: Ciencias; Derecho; Educación	Book - Physical	Book	1	15	2018
Cartoteca: Ciencias; Derecho; Educación; Humanidades; Medicina; Política	Book - Physical	Book	1	11	2018

## 6. Objetos de Analytics

- [Configuración de los objetos de Analytics](#)
- [Suscripción de objetos](#)

Analytics ofrece diferentes herramientas específicas que sirven para presentar, organizar y distribuir la información de Alma. La información de Analytics puede ser visualizada y compartida a través de cinco tipos de objetos:

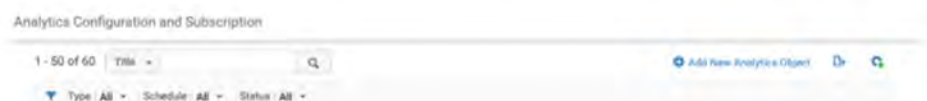
- **Widgets:** pequeñas aplicaciones cuyo objetivo es facilitar la información visual y el acceso a las funciones que se utilizan de forma frecuente.
- **Informes:** herramienta utilizada para explorar los datos y extraer la información significativa.
- **Dashboard o panel de control:** es una herramienta de gestión de la información que reúne colecciones de informes, facilitando el filtrado de datos de acuerdo a determinados parámetros.
- **Scheduled Reports:** informes programados que pueden distribuirse a suscriptores con periodicidad diaria, semanal o mensual.
- **Scheduled Dashboards:** paneles de control programados que pueden ser enviados a suscriptores.

## Configuración de los objetos de Analytics

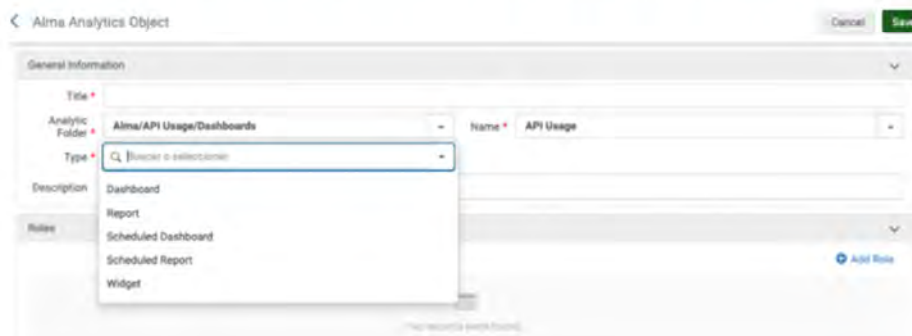
Desde el menú de Analytics seleccione '**Analytics Object List**'.

Analytics Objects  
Analytics Objects List  
★ Analytics User Statistics  
Analytics Citation Attribute Types  
  
Reports  
ANALYTICS. PANEL DE CONTROL  
Portal de Analytics

La página de inicio ofrece una caja de búsqueda y el link que permite añadir un nuevo objeto.



--La primera selección que se deberá realizar es determinar el tipo de objeto que se va a parametrizar.



--Otros campos que deberán ser completados son:

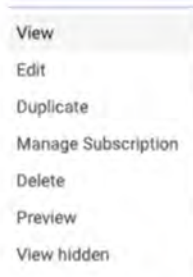
- Título del objeto
- Carpeta de Analytics. Se ofrece un listado con las carpetas del directorio de Alma y el de la propia universidad.
- Nombre del informe incluido en la carpeta de Analytics.
- Descripción: campo no obligatorio, que permite definir el contenido del informe.
- Cuando los informes y paneles de control contienen programación se puede seleccionar entre:
  - Estado: activo o inactivo
  - Programación: diaria, un día de la semana, un día del mes
- Los informes programados pueden ser exportados en formato: excel, pdf y texto.

--Asignar roles. Los informes pueden ser vistos y diseminados a grupos de usuarios con roles específicos.

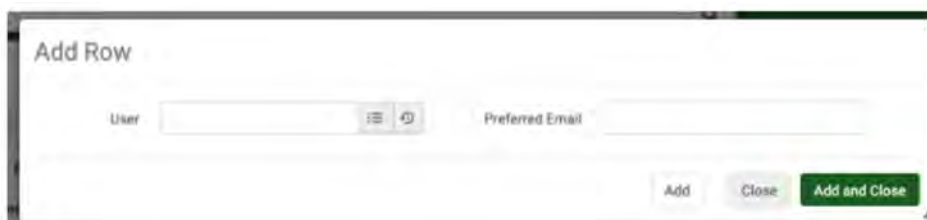
--Finalmente, se guardará la información.

## Suscripción de objetos

Los objetos pueden ser suscritos individualmente desde la opción 'Manage suscription'.



-Añada un fila (raw), y busque en el directorio el nombre del usuario. Cierre y salve la información. El usuario recibirá periódicamente la información parametrizada.

A dialog box titled 'Add Row' with a light gray background and a thin border. It contains two input fields: 'User' and 'Preferred Email'. The 'User' field has a search icon (magnifying glass) to its right. Below the input fields are three buttons: 'Add', 'Close', and 'Add and Close'. The 'Add and Close' button is highlighted in green.



## Visualización de datos (DV)

1. Introducción a la Visualización de Datos
2. Proyecto -1 (Estadística descriptiva)
3. Proyecto -2 (Estadística inferencial)
4. Creación de proyectos con datos de diferentes áreas



# Introducción

- 1. Introducción
- 2. Acceso al módulo de visualización de datos
- 3. Descripción de la página inicial
  - a) Menú principal
  - b) Perfil personal / desconexión
  - c) Crear
  - d) Desplegable de accesos directos :
  - d) Banner central
  - e) Búsqueda de proyectos, paneles de control, informes y datasets
- 4. Áreas temáticas

Last updated [just a moment ago](#)

6 min read

## 1. Introducción

La versión actual de Analytics OAS-DV (Octubre 2020), está construida sobre Oracle (OBIEE 12) y utiliza el lenguaje de consulta estructurada SQL (Structure Query Language). El portfolio de Analytics está formado por 'Oracle Analytics Server' (OAS) y la herramienta independiente 'Data visualization' (DV). La tecnología de OAS se sustenta en la analítica aumentada, cuyas bases son la inteligencia artificial (IA), machine learning (ML) y el procesamiento del lenguaje natural (NLP).

La herramienta 'Visualización de datos' describe de forma gráfica una representación de los datos estadísticos de forma automatizada facilitando:

- una exploración rápida y eficiente de gran cantidad de datos
- detección de modelos estadísticos, correlaciones y tendencias
- proyectos que ayuden a descubrir información sobre cuestiones importantes de la biblioteca

## 2. Acceso al módulo de visualización de datos

El acceso al módulo de **visualización de datos** se realiza desde el menu de Alma Analytics.



## 3. Descripción de la página inicial



El diseño de la página de inicio está estructurada en tres secciones:

### a) Menú principal

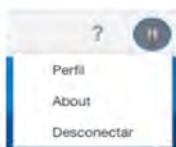
En la parte superior izquierda se localiza el menú principal, cuyo icono es un desplegable. Contiene:

- **Catálogo.** Enlace que facilita un acceso rápido a los informes de Alma-Analytics del área restringida 'Mis carpetas', y a las 'Carpetas compartidas'.
- **Datos.** Enlace que da acceso a lista de las áreas temáticas (Alma-Analytics) y fuentes de datos externas en formato excel.
- **Aprendizaje automático (ML, machine learning).** Es una rama de inteligencia artificial (IA, intelligence artificial) que se centra en desarrollar sistemas que aprenden o mejoran el rendimiento utilizando algoritmos en función de los datos que consumen.
- **Consola.** Sección reservada a la configuración y administración del programa.
- **Vista rápida.** Se trata de una sección de ayuda que contiene tutoriales sobre análisis de datos, agregación de hojas de cálculo, flujos de datos y paneles de control.



### b) Perfil personal / desconexión

En la parte superior derecha se localiza el perfil personal, cambio de contraseña y desconexión.



### c) Crear

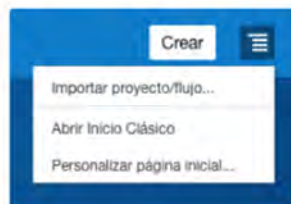
En la sección intermedia se localizan los enlaces a:

- Proyectos**: son visualizaciones que ayudan a comprender datos de forma efectiva.
- Juegos de datos o datasets**: son colecciones de datos tabulados.
- Flujo de datos**: es una herramienta que permite cargar e integrar flujos de datos (stream) en el sistema provenientes de una amplia variedad de fuentes, bases de datos y archivos, para producir conjuntos de datos que puedan ser analizados y visualizados.
- Secuencia**: es una lista de flujos de datos especificados que es útil cuando se desea ejecutar varios flujos de datos como una sola transacción.
- Conexión**: creación de conexiones a bases de datos



### d) Desplegable de accesos directos :

- Importar proyecto/flujo
- Abrir inicio clásico de 'Visualización de datos'
- Personalizar la página inicial



### d) Banner central

Espacio destinado a videos, tutoriales e información de Oracle Analytics. Ofrece acceso a la sección 'Juego de datos o dataset'. Si lo desea esta sección puede ocultarse.



#### e) Búsqueda de proyectos, paneles de control, informes y datasets

Sección final donde se localiza una caja de búsqueda que permite el acceso a los tipos de contenido: *proyectos, paneles de gestión, informes y datasets*. Haga click en un tipo de contenido para agregarlo a la caja búsqueda.



Los contenidos se pueden ordenar por las categorías: nombre, tipo, modificado y propietario.



## 4. Áreas temáticas

Las áreas temáticas son los componentes básicos de la visualización de datos. Una visualización se crea a partir de la elección de un área temática que contenga la información que responda un conjunto específico de cuestiones que quieren analizar.

Técnicamente, las áreas temáticas son agrupaciones de información, denominadas también '**objetos de datos**', que se relacionan entre sí en un contexto particular. Los datos de cada área temática proceden de de los datos extraídos de Alma Analytics.

Las áreas temáticas de 'Visualización de datos' son:



Áreas temáticas	Contenido
AP Usage	Monitoriza el uso de APPs
Analytics objects	Ofrece datos sobre la lista objetos, widgets e informes programados en Analytics.
Analytics Usage Tracking	Facilita el seguimiento del uso de Analytics con detalles de las consultas, sobre las fechas de las consultas y áreas temáticas más usadas.
Benchmark	Facilita la comparación de indicadores de rendimiento entre instituciones similares a través KPIs (Key Performance Indicators) de inventarios, adquisiciones, préstamos, reservas, informes COUNTER.
Borrowing Requests	Información sobre las peticiones de demanda de préstamo interbibliotecario.
Course Reserves	Reservas bibliográficas para cursos. No aplica
Digital Inventory	Información sobre el inventario digital: número de ficheros digitales, objetos digitales, número de descargas, así como detalles de los ficheros, objetos.
Digital Usage	Número de visitas, descargas o fallos de acceso a los objetos digitales, así como detalles de los objetos o fechas de consulta.
E-Inventory	Informes y dashboards de inventarios de portfolios y colecciones electrónicas.
Events	Es un área genérica que investiga eventos generados en Alma. Por ejemplo, presenta actividades de catalogación como la creación de registros, borrado y actualización de MMS-IDs.
Fines and fees	Información sobre sanciones por retrasos o por otras políticas de sanción que aplique la biblioteca. Ofrece una lista de usuarios multados
Fulfillment	Informes y dashboards sobre el área de préstamo.
Funds Expenditure	Información económica sobre fondos asignados, gasto, impuestos aplicados. Detalla información sobre los impuestos aplicados.
Leganto	Es la aplicación de ExLbris orientada a gestionar las listas de bibliografía recomendada. Esta funcionalidad no ha sido contratada por el Consorcio Madroño.
Lending Requets	Información sobre el suministro de préstamo interbibliotecario a otras bibliotecas.
Licenses	Información sobre las licencias de los recursos electrónicos. Responde a preguntas como el número de portfolios y recursos por licencia y los gastos totales por licencias y líneas de pedidos para años fiscales determinados.

<a href="#">Link Resolver Usage</a>	Ofrece datos sobre el resolutor de enlaces OpenURL y una amplia perspectiva sobre cómo funciona este estándar en Alma. Esta información es útil para analizar el desarrollo de la colección electrónica. Ofrece datos de usabilidad de los recursos electrónicos.
<a href="#">Physical items</a>	Ofrece una información detallada del inventario de fondos en papel: títulos, localizaciones, holdings, última fecha de préstamo y transacciones por compra.
<a href="#">Physical Items Historical Events</a>	Ofrece información de los eventos históricos de los ítems físicos y puede responder a cuestiones como el tránsito de los ítems, localizaciones temporales o cambio del estatus de los ítems.
<a href="#">Purchase Requests</a>	Informe que recopila las solicitudes de compra de los usuarios que usan Primo para que las bibliotecas compren un recurso.
<a href="#">Requests</a>	Suministra información de las reservas de los ejemplares físicos: reservas activas, expiradas, solicitudes de reservas que tienen más demanda, reservas por patrón de préstamo o por tipo de material.
<a href="#">Titles</a>	Ofrece información sobre el número de títulos de ítems físicos, holdings, portfolios, número de colecciones electrónicas. Suministra medidas del uso de los recursos bibliográficos o cómo vincular los holdings físicos con los portfolios para disponer de una representación de la colección.
<a href="#">Usage Data</a>	Suministra datos de los informes COUNTER: (Release 5): DR (Database Reports), PR (Platform Reports), TR-B (Title Master Requests -Books) TR-J (Title Master Request-Journals).
<a href="#">Users</a>	Facilita información detallada de los usuarios como roles, categorías, tipologías, grupos de usuarios y datos demográficos.

# Proyecto - 1

- 1. Fase inicial
- 2. Etapas en la creación de un proyecto
  - 2.1. Escenario
  - 2.2. Crear un proyecto
  - 2.3. Agregar juego de datos
  - 2.4. Guardar el proyecto
  - 2.5. Selección de datos
  - 2.6. Filtrar los datos
    - a) Banner superior del lienzo
    - b) Funciones SQL
  - 2.7. Tipos de visualización
  - 2.8. Proyecto. Resultado

## 1. Fase inicial

Un proyecto es una representación gráfica que tiene como finalidad analizar los datos de forma eficiente, de manera que sirva para la toma de decisiones fundamentadas.

Un proyecto de visualización de datos implica:

- a) definición del escenario de estudio
- b) selección y análisis de los datos
- c) selección del tipo de gráficos



## 2. Etapas en la creación de un proyecto

### 2.1. Escenario

Consiste en una breve narración que describa en el planteamiento del proyecto desde un enfoque estructurado que implique una combinación de varios elementos: *variables que intervienen, descripción de los datos y el soporte visual utilizado.*

Esta información se describe en el apartado '**Narrar**'.

El proyecto que se describe a continuación visualiza los datos de préstamos bibliográficos a estudiantes de grado.



### 2.2. Crear un proyecto

Pulse en dicho botón **Proyecto**. A continuación, se mostrará el cuadro de diálogo

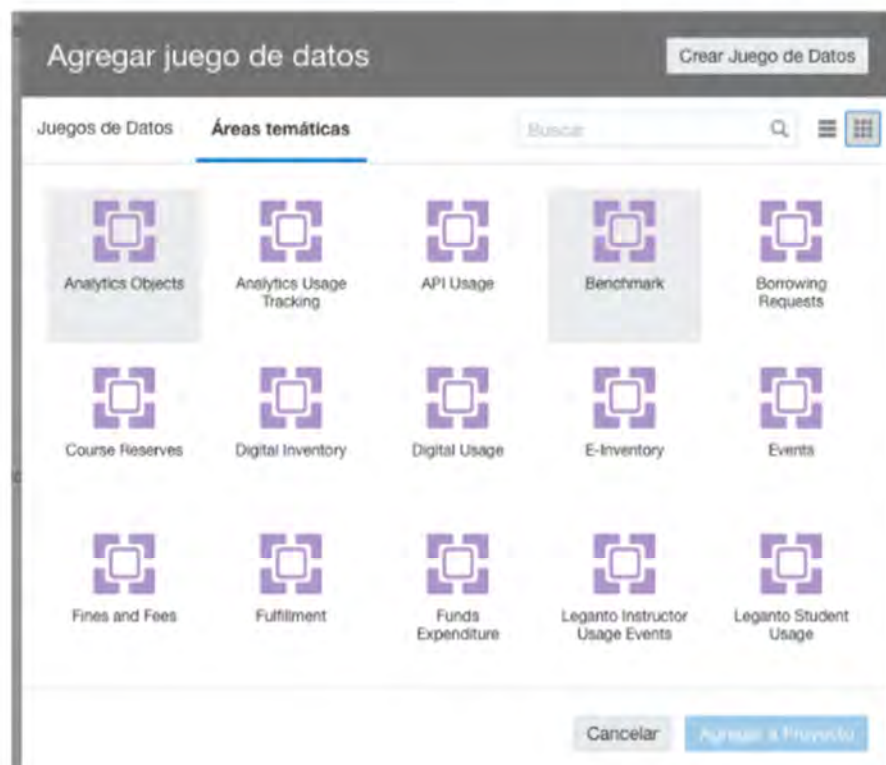
**Agregar juego de datos o Dataset**

### 2.3. Agregar juego de datos

Este cuadro remite al origen de los datos del proyecto, cuyas fuentes son:

- a) Juego de datos o datasets. Son ficheros de datos tabulados en excel, importados de fuentes externas a Alma-Analytics.
- b) **Áreas temáticas** de Alma-Analytics.

Si no encuentra el origen de los datos puede hacer click en '**Crear Juego de datos**'.



## 2.5. Selección de datos

Para construir la visualización, se seleccionarán las variables correspondientes.

Estadísticamente, las **variables de clasificación** en cuantitativas y categóricas.  
Este programa diferencia las variables entre:

-- **cuantitativas** con el símbolo # (almohadilla o numeral)



-- **categóricas**, representadas por el símbolo A (acrónimo de 'Atributo')



-- **fecha**, variable en escala que permite calcular intervalos, puesto que tiene sentido calcular la diferencia entre valores. Está representada por un reloj



## 2.4. Guardar el proyecto

El programa ofrece las siguientes opciones para salvar los proyectos: *guardar*, *guardar como* y *guardar automáticamente*.

Para minimizar el riesgo de pérdida de datos, se recomienda guardar el proyecto con un nombre, y a continuación seleccionar '*guardar*' o '*guardar automáticamente*'.

Por defecto, el sistema guardará los proyectos en el área privada de Alma-Analytics: **Mis carpetas**.






## 2.6. Filtrar los datos

Hay diferentes procedimientos para filtrar los datos:

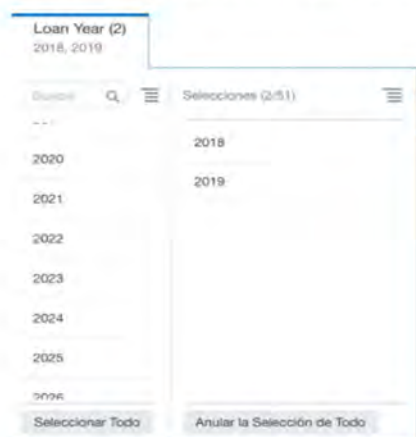
### a) Banner superior del lienzo

El margen superior del lienzo contiene un espacio con el signo  'Haga clic aquí o arrastre datos para agregar un filtro', donde las variables pueden ser filtradas.



Por ejemplo:

--Filtro de los años de préstamo, se pulsa en la variable y añade los años del proyecto



### b) Funciones SQL

Para ampliar información sobre el lenguaje SQL consulte la [tabla de sintaxis SQL](#). Por ejemplo:

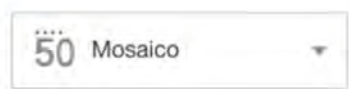
Función: tiempo	Sintaxis	Dexcripción
TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Return date = TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Items devueltos el día 7 antes de la fecha actual.
	Return date <= TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Items devueltos con anterioridad a 7 días respecto a la fecha actual
	Return date >= TIMESTAMPADD(SQL_TSI_DAY, -7, CURRENT_DATE)	Items devueltos entre hace 7 días y la fecha actual.



2.7. Tipos de visualización

La selección de las gráficas se puede realizar desde el banner intermedio '50 Mosaico' en el margen izquierdo, o desde el **icono de gráfico de barras**, margen derecho.

Para seleccionar la categoría de gráfico que se adecua al análisis puede consultar: '[Representaciones gráficas](#)'.



Listado completo de gráficos



## Proyecto préstamo domiciliario por categoría de usuarios. Año anterior



## Proyecto -2

- 1. Escenario
  - Regresión lineal y línea de tendencia
- 2. Fases del proyecto
  - 2.1. Crear un proyecto
  - 2.2. Agregar las áreas temáticas
  - 2.3. Guardar el proyecto
  - 2.4. Selección de datos
  - 2.5. Filtrar los datos
  - 2.6. Selección gráfico de dispersión
  - 2.7 Proyecto. Resultado

## 1. Escenario

### Regresión lineal y línea de tendencia

Se analiza la asociación entre las variables cuantitativas, número de préstamos domiciliarios y el número de ítems en el repositorio en los 5 últimos años. Se busca el mejor modelo que explique la relación entre ambas variables y genere una previsión.

Se considera como variable dependiente el número de préstamos y como variable independiente el número de ítems.

## 2. Fases del proyecto

### 2.1. Crear un proyecto

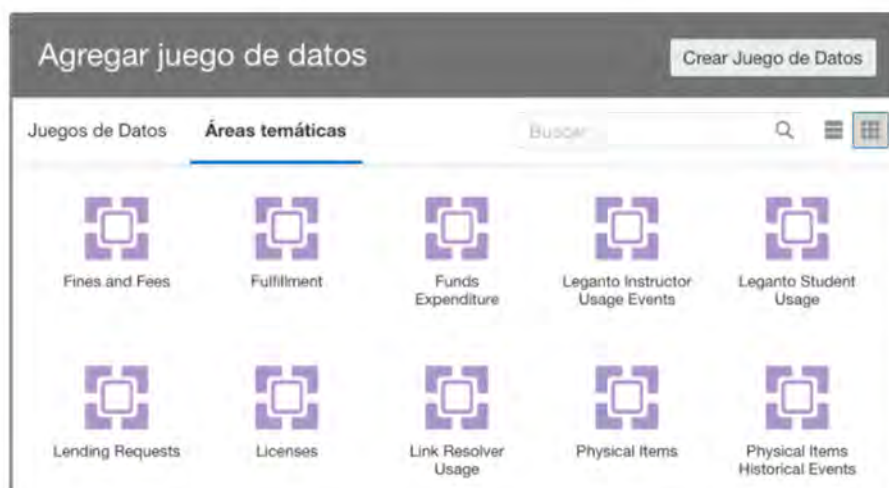
Pulse en dicho botón **Proyecto**. A continuación, se mostrará el cuadro de diálogo **Agregar juego de datos o Dataset**

### 2.2. Agregar las áreas temáticas

Se selecciona el área temática 'Physical items'

### 2.2. Agregar las áreas temáticas

Se selecciona el área temática 'Physical items'





### 2.3. Guardar el proyecto

Se recomienda guardar con un nombre el proyecto en el área privada 'Mis carpetas'.

### 2.4. Selección de datos

(a) Área temática 'Physical items'

Utilice para la visualización de datos el formato 'Tabla'



Arrastre el lienzo las dos variables cuantitativas:

-->Number of loans (not in house)

-->Number of items (in repository)

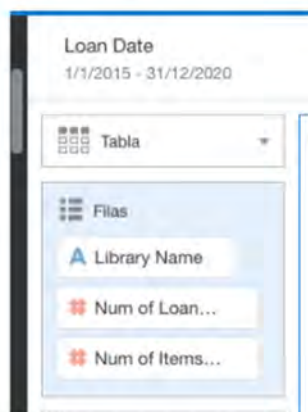
Arrastre al lienzo la variable cualitativa:

--> Library name

### 2.5. Filtrar los datos

Se filtran los datos por los 5 últimos años de préstamo (2015-2020). Para ello se seleccione el área temática 'Fulfillment'.

Arrastre el criterio 'Loan date' al banner superior del lienzo y seleccione las fechas.

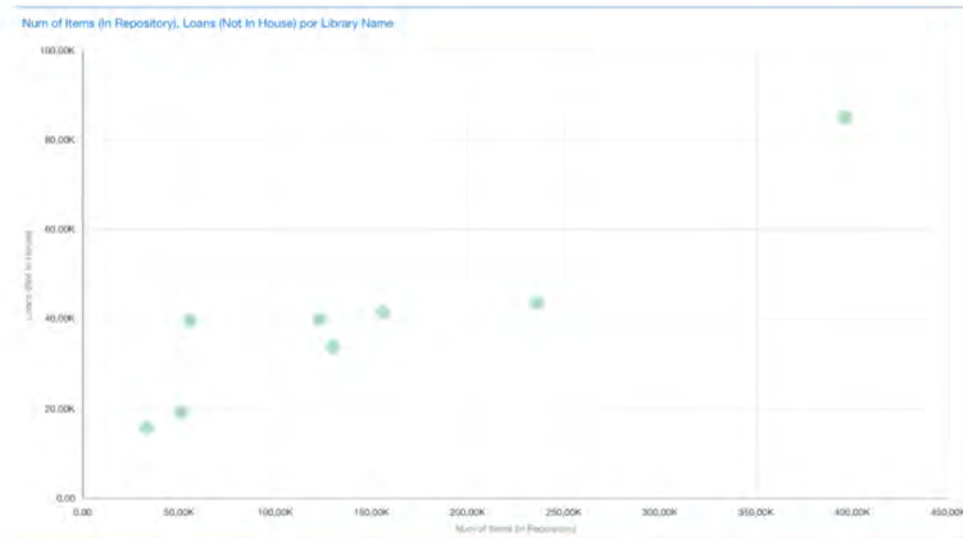


## 2.6. Selección gráfico de dispersión

Seleccione el gráfico de dispersión



Automáticamente se creará una gráfica de nube de puntos con la relación entre los préstamos y el número de ítems en el repositorio.



A continuación, añada la línea de tendencia. Para ello, pulse con el botón derecho en el lienzo y aparecerá una ventana que entre otras opciones incluye la 'línea de tendencia'.

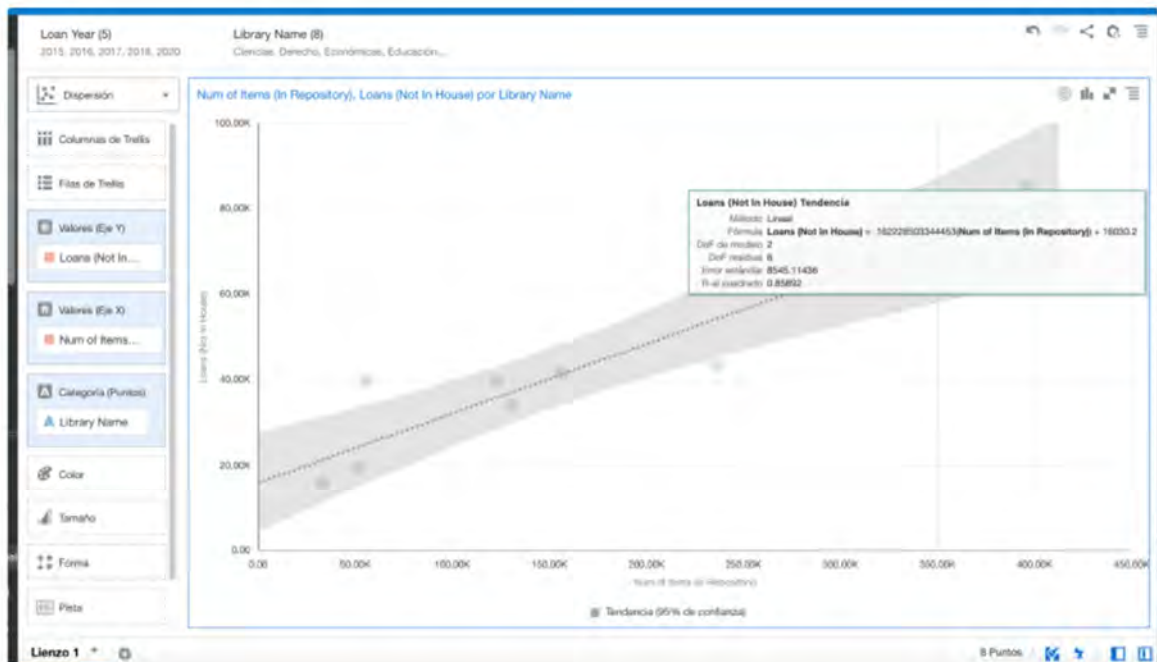
A continuación, añada la línea de tendencia. Para ello, pulse con el botón derecho en el lienzo y aparecerá una ventana que entre otras opciones incluye la 'línea de tendencia'.



## 2.7 Proyecto. Resultado

Se puede considerar un modelo bueno. Es decir, la variabilidad del préstamo estaría explicada en un 85,89% por el número de ítems en el repositorio.

$$R^2 = 85,89$$



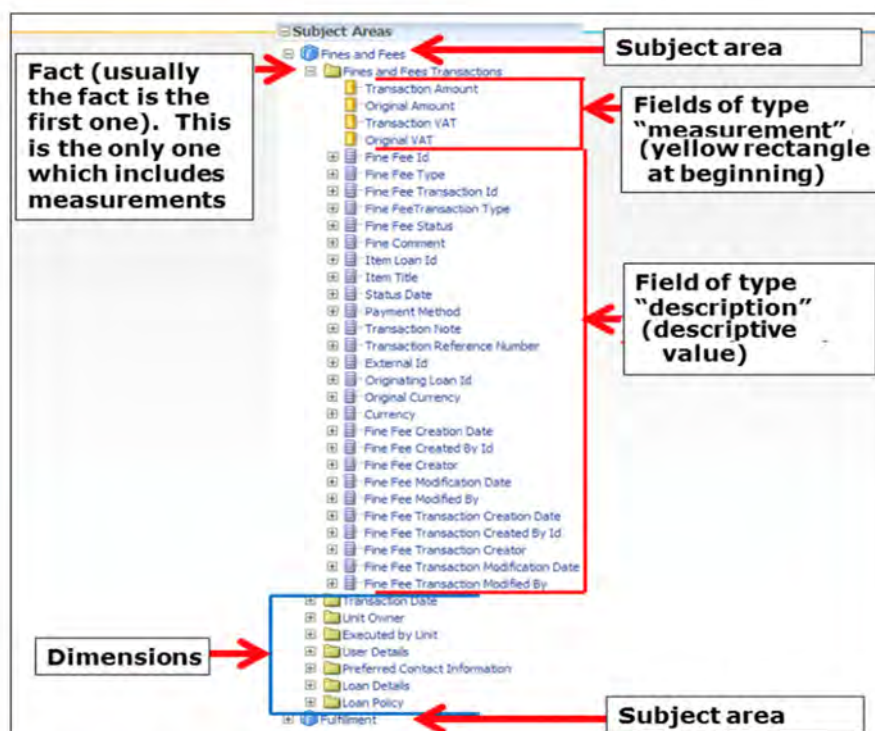
## Creación de proyectos con datos de diferentes áreas

- 1. Escenario
- 2. Fases del proyecto
  - 2.1. Exportación datos a Excel
  - 2.2. Visualización de datos. Juego de datos o dataset
  - 2.3. Cargar el fichero Excel
  - 2.4. Agregar los datos
  - 2.5. Crear un proyecto
  - 2.6. Guardar el proyecto
  - 2.7 Añadir un área temática
  - 2.8 Agregar el área temática 'Physical items'
  - 2.9 Preparar el proyecto
  - 2.10 Diagrama de datos
  - 2.11 Visualizar los datos
  - 2.13. Resultado

'Visualización de datos' permite crear informes avanzados combinando datos de más de un área temática. Se trata de un proceso complejo que requiere seguir las reglas de dependencias entre datos.

Un área temática contiene los siguientes elementos:

- (a) **Una tabla de datos (Fact table):** generalmente se sitúa al inicio del área temática. Comprende las mediciones, métricas o datos que forman la parte central del área temática.
- (b) **Dimensiones (Dimensions):** son atributos o campos descriptivos textuales o números discretos que se comportan como texto. Los atributos de dimensión como el código de biblioteca, el nombre del proveedor o la fecha de préstamo contextualizan los datos.
- (c) **Columnas de medida (Measurement Columns):** contienen listas de valores de datos que pueden contarse o agregarse.
- (d) **Columnas descriptivas (Descriptive columns)** son listas de valores categóricos. No existe una relación jerárquica entre estos valores.
- (e) **Columnas jerárquicas (Hierarchical columns)** son valores que se organizan utilizando relaciones padre-hijo. Esta columna presenta una estructura en forma de árbol.



Fuente: Exlibris General Terminology. URL:

[https://knowledge.exlibrisgroup.com/Alma/Product\\_Documentation/010Alma\\_Online\\_Help\\_\(English\)/080Analytics/010Introduction/General\\_Terminology](https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_(English)/080Analytics/010Introduction/General_Terminology)

Requerimientos para crear proyectos de diferentes áreas temáticas:

- Debe haber al menos **una dimensión común** entre las dos áreas temáticas.
- Los **campos descriptivos deben proceder de dimensiones comunes**.



## 1. Escenario

El objetivo del proyecto es conocer cuántas peticiones de compra realizadas por los usuarios en Primo están disponibles en préstamo, del total de peticiones realizadas. Es decir, se trata de determinar el número de peticiones que disponen del campo 'Permanent call number' y son susceptibles de ser prestadas.

La posibilidad de unir firmas a áreas temáticas como 'Funds expenditure' o 'Purchase requests' era una de las limitaciones de los informes de Alma-Analytics y que 'Visualización de datos' ha corregido.

Este proyecto contiene las peticiones que han sido valoradas y aprobadas por la biblioteca. Los pedidos han sido tramitados por 'Adquisiciones' y el estatus de las POLs es 'cerrada'. De acuerdo con el workflow de la biblioteca la siguiente etapa consistiría en completar la firma y el registro de ejemplar de los ejemplares recibidos.

## 2. Fases del proyecto

### 2.1. Exportación datos a Excel

Se exporta a un fichero Excel las peticiones solicitadas por los usuarios en 2020 que cumplen la condición 'Aceptadas' y cuya factura está 'cerrada'. El fichero contiene 748 títulos.

Formato de exportación:

MMS Id	Purchase Request Title	PO Line Reference	POL Status	Peticionario Primary Identifier
991007218279704211	Manual de derecho civil	POL-19777	CLOSED	51780046P
991007883255004211	Productos financieros y seguro	POL-19662	CLOSED	46833377H
991007885656604211	Human Rights between Law and Politics: The Margin of Appreciation in Post-national Contexts	POL-19764	CLOSED	05220048Z
991007885656704211	General Principles of Law: European and Comparative Perspectives	POL-19766	CLOSED	05220048Z
991007885656804211	: Hablamos de Derechos Humanos: La doctrina del TEDH y su aplicación en España desde los votos particulares del Juez Paulo Pinto de Albuquerque	POL-19771	CLOSED	05220048Z

### 2.2. Visualización de datos. Juego de datos o dataset

Desde 'Crear' se pulsa en 'Juego de datos'.



### 2.3. Cargar el fichero Excel

A continuación, **arrastre** el fichero Excel al espacio indicado en la imagen.



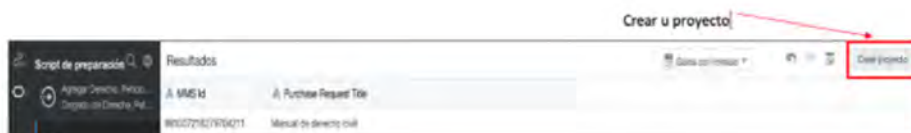
### 2.4. Agregar los datos

Agregue los datos Excel a 'Visualización de Datos' pulsando en el botón **'Agregar'**.



### 2.5. Crear un proyecto

Los campos del fichero Excel y el área temática de Analytics deben definirse de la misma manera en 'Visualización de Datos'. En este ejemplo, los dos campos (MMS-ID, Título de la petición de compra) como atributos o campos descriptivos textuales. Haga clic en el botón **'Crear proyecto'** para crear el conjunto de datos y guardarlo.



### 2.6. Guardar el proyecto

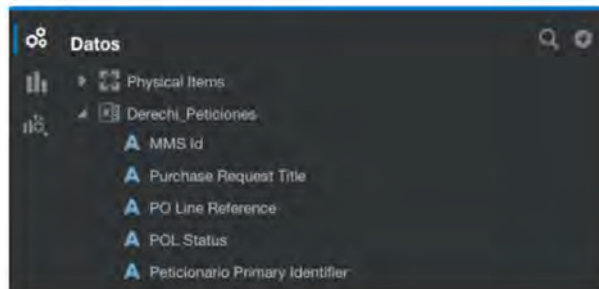
Guarde el proyecto con un nombre en el espacio privado de Alma-Analytics **'Mis carpetas'**.

## 2.6. Guardar el proyecto

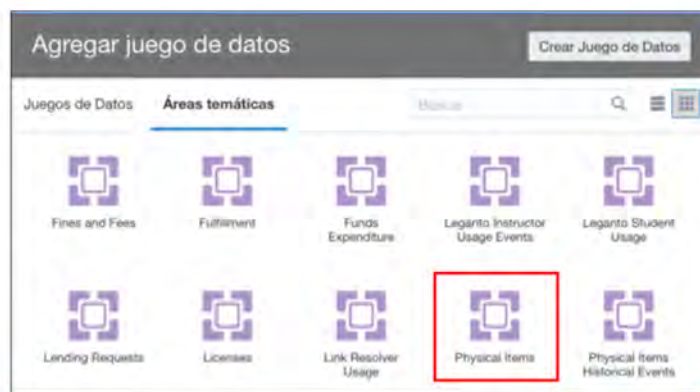
Guarde el proyecto con un nombre en el espacio privado de Alma-Analytics 'Mis carpetas'.

## 2.7 Añadir un área temática

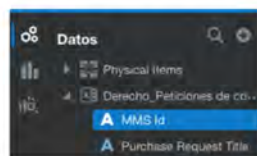
Pulse en el símbolo  para añadir un área temática.



## 2.8 Agregar el área temática 'Physical items'

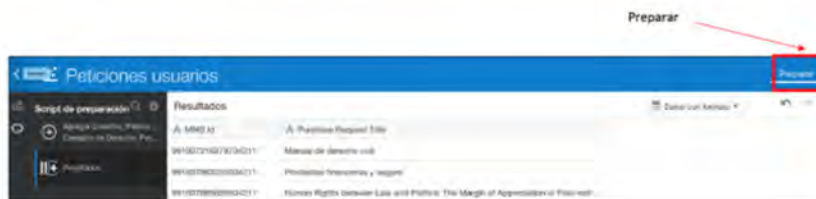


El área temática 'Physical items' se sitúa en la parte superior de la lista.



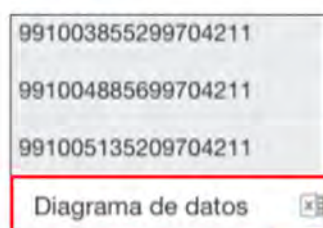
## 2.9 Preparar el proyecto

Pulse en el botón '**Preparar**' en la parte superior del lienzo.



## 2.10 Diagrama de datos

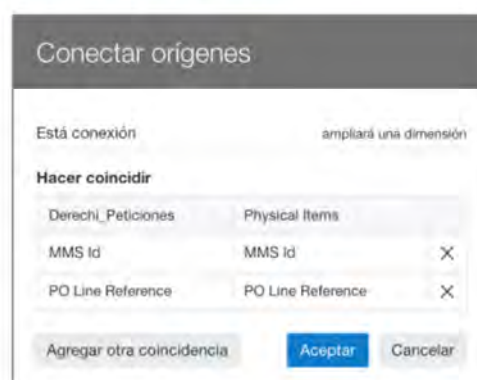
Al final del lienzo pulse en '**Diagrama de datos**'.



Pulse en el número '**2**', que indica que hay un campo de conexión entre las dos fuentes de datos.



La **dimensión común** entre las dos áreas temáticas '**Purchase requests**' y '**Physical items**' se realiza a partir del '**Título**'. A continuación, se pulsa en '**Aceptar**'.



## 2.11 Visualizar los datos

Pulse en el botón 'Visualizar'



## 2.12. Añadir el campos y filtros

--> Descarga ficheros Excel en 'Visualización de datos' :

Conserve los encabezamientos Excel, y elimine la celdas combinadas.

(a) Añadir criterios:

--Dataset (Excel): Purchase Request (Title)

--Dataset (Excel) : MMS-ID

--Dataset (Excel) : POL Status >Closed

--Dataset (Excel): POL Line Reference

--Physical Items : Holding Details : Permanent call number

Al añadir del 'Área temática' el criterio 'Permanent call number' carga todas las signaturas, asociadas y no asociadas a los criterios del Dataset.

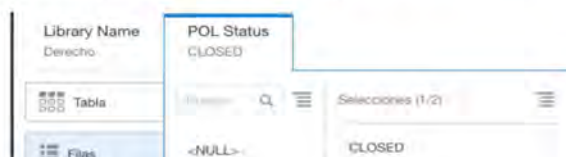
(b) Por ello, se aconseja crear un filtro previo por biblioteca.

(c) En esta fase ya están asociadas las signaturas con los títulos. Sin embargo, si se exporta el proyecto se puede comprobar que se han añadido todas las signaturas de la biblioteca seleccionada, cuyos títulos no asociados a las signaturas están vacíos, POL Line Reference ( -1 ) y POL status (NULL).

Purchase Request Title	Permanent Call Num	PO Line Reference	POL Status
Wissenschaft, Praxis und Dogmat D/Oa/584		POL-21913	CLOSED
XConcepto de justicia ante el retc D/Ig.1/691		POL-20783	CLOSED
Y de nuevo, la eutanasia : D/Ig.2/555		POL-22560	CLOSED
extraterritoriales et droit intern D/PEDIDO/IPB/41		POL-23309	CLOSED
¿Derecho penal liberal o Derecho D/Qa.2.2/157		POL-22491	CLOSED
¿Se convertirá la sentencia del Jul D/PEDIDO/HIS/54		POL-22751	CLOSED
Ética judicial D/Pd/683		POL-21223	CLOSED
	ADM/		-1
	ADM/		-1
	ADM/		-1
	ADM/		-1

Captura de pantalla 2020-12-13 a las 13.20.51.png

(d) Para eliminar las signaturas no asociadas al dataset se arrastra al banner superior una segunda copia del dataset 'POL Status' y se filtran los datos por 'CLOSED'.





Automáticamente, se obtendrán las firmas de las peticiones compra aprobadas con POL (Status) 'Closed'.

### 2.13. Resultado

- Peticiones de compra solicitadas por los usuarios. POLs status 'Cerrada: 803 títulos
- Títulos con firma: 794 títulos
- Títulos sin firma 'Unknown': 7 títulos
- El resto de los títulos: 9 títulos contienen firmas, pero no figuran en el listado final.
  - >El mensaje 'Alerts' del PO advierte que el registro es parte de una serie. Esto puede significar que se utilizó un tipo de línea de orden de compra de monografía para un registro bibliográfico de revista / serie.
    - "Items already exist in the repository, The bib record is part of a series/book set"
    - "Items already exist in the repository, Duplicate active orders, The bib record is part of a series/book set."
- >El estatus PO es (sent).

# Análisis estadísticos

1. [Análisis de clúster](#)
2. [Series temporales y modelos ARIMA](#)
3. [Regresión lineal simple y línea de tendencias](#)



## Análisis de clúster

- [Introducción](#)
- [Razones para la clasificación](#)
- [Selección de datos](#)
- [Tipos de variables](#)
- [Métodos de agrupación de variables](#)
  - [1. Método jerárquico](#)
    - [Distancias entre conglomerados](#)
  - [2. Método no jerárquico o k-means](#)
    - [K-Means](#)
- [Ejemplo de análisis clúster. Método jerárquico](#)

### Introducción

El análisis clúster es una técnica multivariante cuya idea básica consiste en agrupar elementos (variables) formando conglomerados (clústers) que tengan la máxima homogeneidad interna dentro de cada grupo y la mayor diferencia entre grupos. Es un análisis descriptivo, no inferencial, que se centra en clasificar los objetos en clases o conglomerados. Es decir, permite detectar un número óptimo de grupos y su composición a partir de la **similitud** existente entre ellos, de forma que los grupos tengan un alto grado de homogeneidad interna y un alto grado de heterogeneidad externa.

Los principales objetivos del análisis de clúster son:

- Clasificar los datos para facilitar su interpretación
- Investigar esquemas conceptuales de agrupación
- Generar hipótesis a través de la exploración de datos

### Razones para la clasificación

La clasificación representa un método eficiente para organizar un amplio conjunto de datos, de modo que la información se pueda entender y recuperar más fácilmente. La necesidad de resumir conjuntos de datos mediante la agrupación basada en la similitud puede proporcionar una descripción muy eficiente para detectar patrones y diferencias entre grupos.

El análisis de clúster trata de explorar conjuntos de datos para evaluar si pueden o no, de manera significativa formar grupos que se parezcan entre sí y que sean distintos a su vez en algunos aspectos de otros conglomerados.

El análisis de clúster es una técnica eminentemente exploratoria que no ofrece soluciones únicas, las soluciones dependen de las variables usadas en el estudio para revelar las características inherentes al grupo y del método utilizado. En la clasificación de datos, además de los criterios estadísticos existen factores subjetivos que pueden influir en las clasificaciones como son, el objeto de estudio, el conocimiento previo de los datos y la decisión sobre el número óptimo de clústers.

### Selección de datos

La selección de las variables es un factor clave en la formulación del análisis de clúster. El conjunto de variables debe describir la similitud de los objetos en términos relevantes. Como enfoque general, se puede decir que un mayor número de variables no va a conseguir una mayor homogeneidad de los grupos, sino que al contrario puede debilitar su semejanza.

Para realizar un análisis de clúster se necesita una medida de **similitud o semejanza entre los objetos analizados**. Las técnicas de agrupamiento se basan en algoritmos que determinan la semejanza, proximidad y similitud entre las variables.

Un aspecto previo a considerar es si se van a utilizar los datos tal como son obtenidos o se estandarizan cuando las medidas están en escalas distintas o determinados valores tienen una alta capacidad discriminante. Antes de estandarizar las variables se debería determinar si la transformación es necesaria ya que puede conllevar un cambio en las relaciones entre las variables, un cambio de similitud y en consecuencia una variación en el resultado del análisis.

La fórmula más común es estandarizar los datos, es restar a cada observación la media de la variable, el resultado se divide entre la desviación típica. Donde ' $z_i$ ', es la estandarización de la variable, ' $x_i$ ' es cada una de las observaciones, ' $\bar{x}$ ' (barra) es la media y ' $s$ ' desviación típica. En SPSS la estandarización se realiza automáticamente en el análisis de 'clúster jerárquico', método 'estandarizar', puntuaciones ' $z$ '.

### Tipos de variables

- **Variables cuantitativas:** se expresan con un valor numérico y pueden clasificarse en:
  - Variables discretas, no poseen valores intermedios entre ellas. Se expresan con un número, sin decimales. Por ejemplo '5'.
  - Variables continuas, pueden adoptar cualquier intervalo o valores en medio de dos exactos. Generalmente están representados por valores decimales. Por ejemplo, 5,75.
- **Variables categóricas,** también denominadas 'cualitativas'. Son variables que describen cualidades o circunstancias. Se pueden clasificar:
  - Nominales, no siguen ningún orden o lógica. Por ejemplo, colores.
  - Ordinales, son variables con un orden o jerarquía. Por ejemplo. las notas de un examen (suspense, aprobado, sobresaliente).
  - Dicotómicas o binarias, son variables nominales con sólo dos categorías. Por ejemplo, hombre-mujer.

Antes de iniciar el análisis de clúster se deben tomar 3 decisiones:

- **Selección de variables relevantes** para identificar a los grupos de acuerdo al objetivo del estudio.
- **Selección del procedimiento de agrupación de las variables,** jerárquicos o no jerárquico 'k-means'.
- **Número de conglomerados,** que dependerá del resultado de la agrupación de variables.



## Métodos de agrupación de variables

Se distinguen dos métodos de agrupación de las variables:

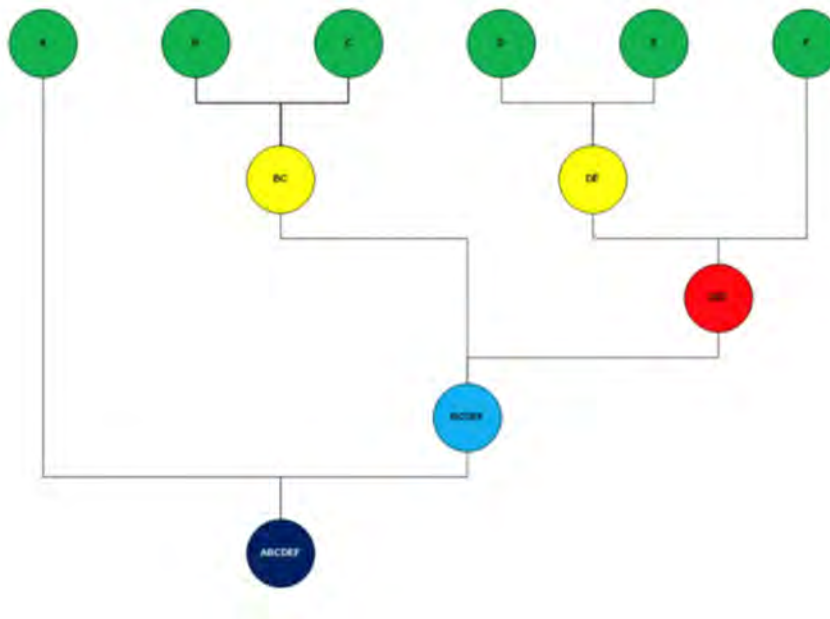
### 1. Método jerárquico

El método jerárquico por su exhaustividad resulta conveniente para análisis de conglomerados con un número de variables no excesivamente grande. Cuando se trata de muestras que contienen un número alto de observaciones se utiliza el método de agrupamiento no jerárquico '**k-means**'.

Los métodos jerárquicos se subdividen en:

**El método jerárquico aglomerativo**, comienza con tantos clusters como elementos tengamos que clasificar y en cada paso se recalculan las distancias entre los grupos existentes, se unen los grupos más similares y se separan los más distantes. El algoritmo acaba con un clúster que contiene todos los elementos. Este método utiliza las técnicas de distancia:

- Método de linkage simple o vecino más cercano
- Método de linkage completo o vecino más lejano
- Método del promedio entre grupos
- Método de la mediana
- Método de Ward

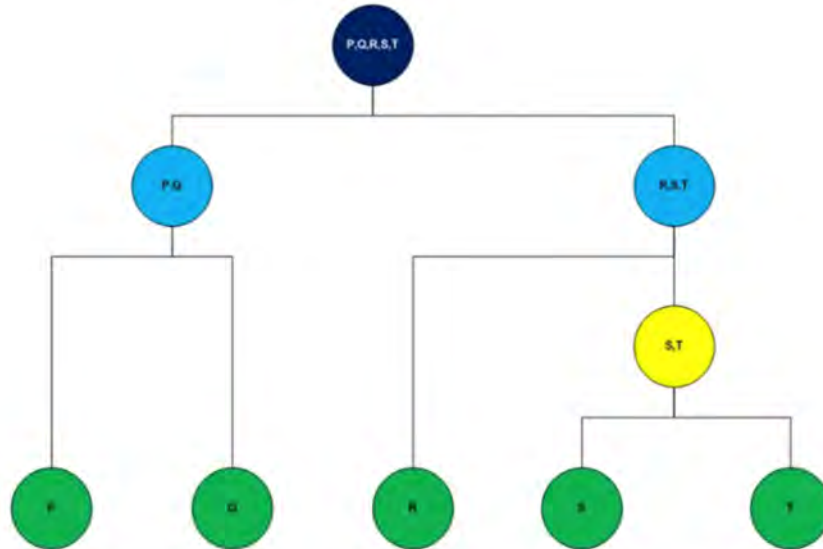


Dendrograma de clúster jerárquico aglomerativo. Fuente: *Clustering jerárquico con Python*  
([https://www.jacobsoft.com.mx/es\\_mx/clustering-jerarquico-con-python/](https://www.jacobsoft.com.mx/es_mx/clustering-jerarquico-con-python/))

**El método jerárquico divisivo**, se parte de un único conglomerado comienza que engloba a todos los elementos y en cada paso se van formando grupos cada vez más diferentes. El algoritmo acaba con tantos clústers como elementos se hayan clasificado.

Este método utiliza las técnicas de distancia:

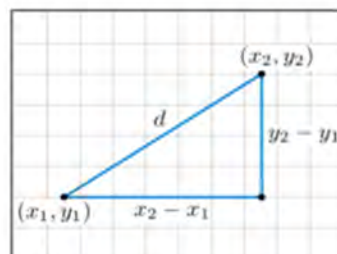
- Método de linkage simple
- Método de linkage completo
- Método de promedio entre grupos
- Método de la mediana
- Método de Ward



**El dendrograma** es el gráfico de representación en forma de árbol utilizado con el método jerárquico que permite observar cómo se han formado los clústers a partir de las subdivisiones o agrupamientos. Esta herramienta ayuda a decidir el número de conglomerados que pueden representar mejor la estructura de los datos.

#### Distancias entre conglomerados

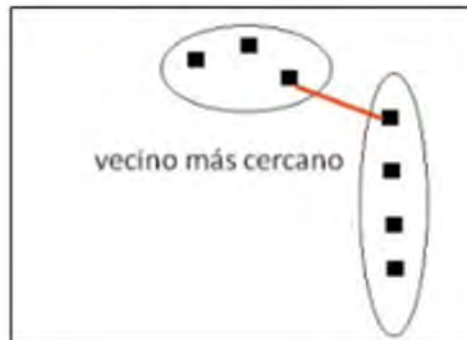
Para ambos casos, la medida que se utiliza es la **distancia euclidiana**, que consiste en hallar la distancia entre dos puntos en un plano cartesiano.



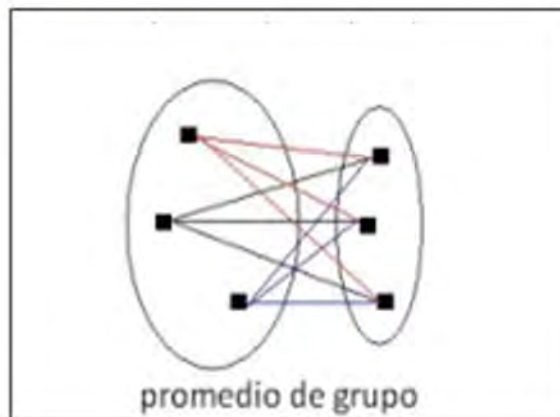
$$d_E(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Dentro de los métodos jerárquicos acumulativos las técnicas para la obtención de conglomerados se subdividen en:

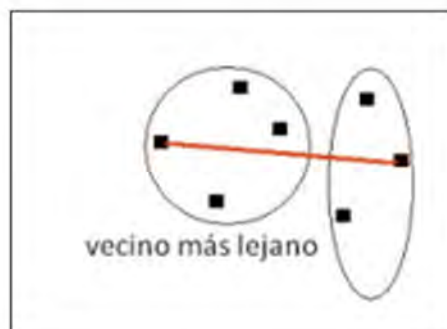
- Vecino más cercano o distancia mínima: los elementos que se combinan en cada grupo son aquellos que tienen una menor distancia o mayor similitud. Posteriormente se recalcula la distancia del clúster respecto al resto de casos formándose la siguiente agrupación mediante el mismo criterio.



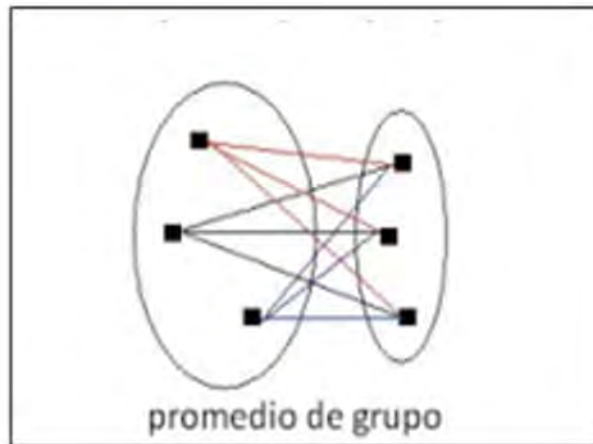
- Promedio de grupos, mide la proximidad entre dos grupos calculando la media de las distancias o similitudes entre elementos de ambos grupos.



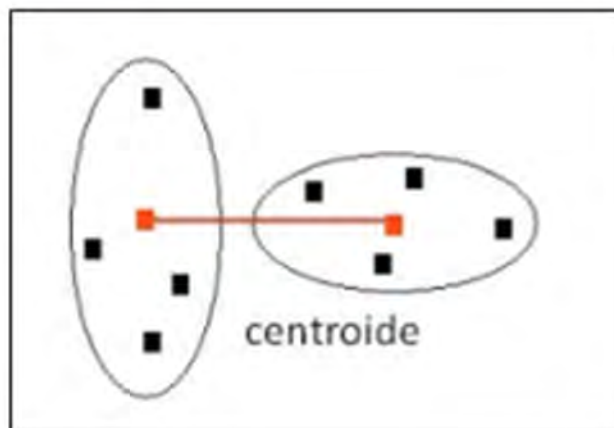
- Vecino más lejano o distancia máxima: la distancia se calcula a partir de la distancia de los dos puntos más alejados.



- Promedio de grupos, mide la proximidad entre dos grupos calculando la media de las distancias o similitudes entre elementos de ambos grupos.



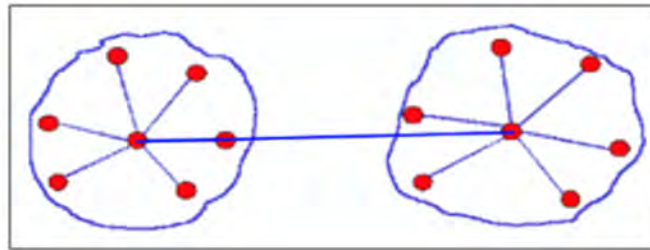
- Método de la mediana, mide la distancia entre los centros de gravedad de dos grupos **centroides**, continúa agrupando conglomerados cuya distancia entre los centroides sea mínima. Con este procedimiento se reducen los casos extremos.



- Método de Ward: este procedimiento intenta formar conglomerados minimizando la varianza interna. Para ello se calculan los centroides de los grupos y las distancias de las observaciones con respecto al centroide (suma de cuadrados). En la última etapa, se agrupan los conglomerados que generan el menor incremento de distancias dentro de cada conglomerado. Este procedimiento crea grupos homogéneos.

/

**i** Analytics Data Visualization crea automáticamente jerarquías de clústers por **el método jerárquico acumulativo o ascendente, o el método divisivo o descendente.**



## 2. Método no jerárquico o k-means

Los métodos no jerárquicos categorizan los elementos **según un número de clústers establecido. Se necesita que el número de particiones esté fijada de antemano.**

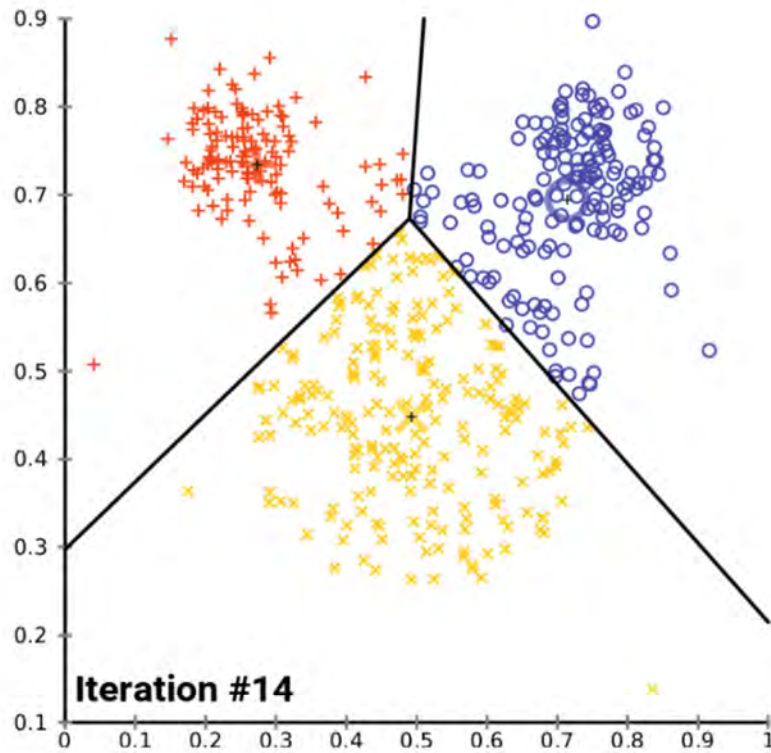
Una clasificación no jerárquica está formada por grupos homogéneos sin formar relaciones entre ellos.

### K-Means

Es un método de agrupamiento utilizado para gran cantidad de datos. Es un método fácil de implementar y muy eficiente que necesita como dato de entrada el número de grupos en los que se van segmentar los elementos. A partir de este número k de clústers, el algoritmo coloca primero k puntos aleatorios (centroides). Luego asigna a cualquiera de esos puntos todas las muestras con las distancias más pequeñas.

A continuación, el punto se desplaza a la media de las muestras más cercanas. Esto generará una nueva asignación de muestras, ya que algunas muestras estarán ahora más cerca de otro centroide. Este proceso se repite de forma iterativa y los grupos se van ajustando hasta que la asignación no cambia más moviendo los puntos. Este resultado final representa el ajuste que maximiza la distancia entre los distintos grupos y minimiza la distancia intragrupo.





Fuente: CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>), from Wikimedia Commons

- i** Analytics Data Visualization especifica que la agrupación en clúster k-means particiona las observaciones en k clusters en los que cada observación pertenece al clúster con la media más próxima.

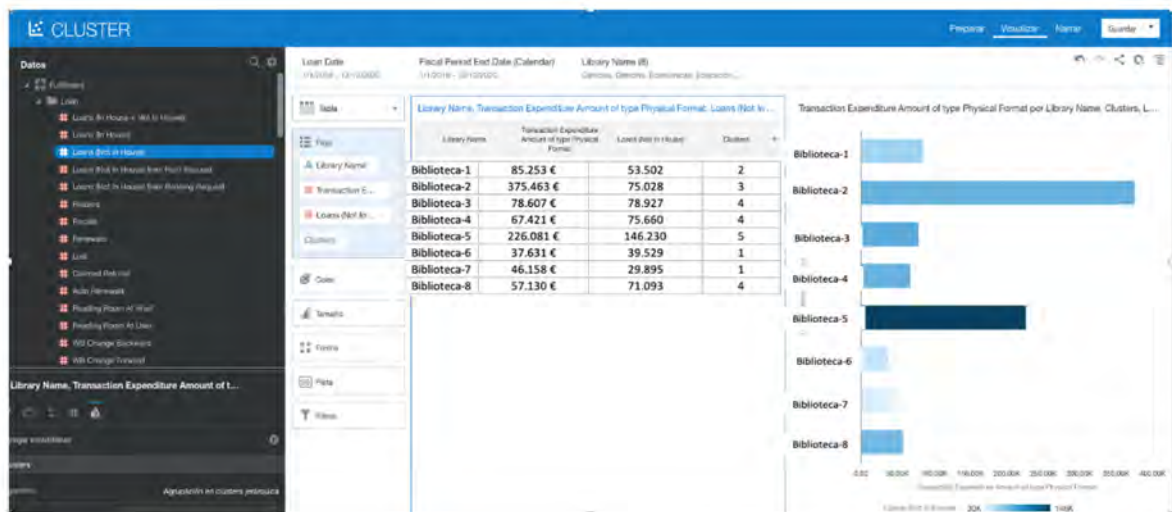
### Ejemplo de análisis clúster. Método jerárquico

Aplicación del análisis clúster, método de agrupación jerárquico para la clasificación de las bibliotecas de una universidad. El objetivo del análisis es clasificar las bibliotecas en función de sus similitudes, es decir obtener los grupos de bibliotecas homogéneas que se parezcan entre sí respecto a características comunes.

Parámetros:

Input: presupuesto bibliotecas, años (2018-2019)

Output: préstamos domiciliarios, años (2018-2019)



# Series temporales y modelos ARIMA

- Series temporales
  - Componentes de una serie temporal
    - Tendencia (T) (Trend)
    - Componente cíclica ( C ) (Cyclic)
    - Componente estacional (S) (Seasonal)
    - Componente aleatorio o irregular (I) (Irregular)
  - Clasificación de las observaciones de las series temporales
    - a) Procesos estacionarios.
    - b) Variable error o ruido blanco
    - c) Procesos estacionales
- Modelos ARIMA
  - Modelos autorregresivos (AR)
  - Modelo de medias móviles (MA)
  - Modelo Integrado (I)
  - Modelos ARMA
  - Condiciones del modelo ARIMA
- Modelos SARIMA o ARIMA con componente estacional
  - Referencias:

## Series temporales

Una serie temporal es una secuencia ordenada de valores que una variable toma a lo largo del tiempo, donde cada una de las observaciones está asociada a un momento de tiempo. Cada observación se denota como ' $Y_t$ '. Las series temporales posibilitan observar la evolución de una variable a lo largo del tiempo, analizar su dinámica, y predecir el comportamiento en el futuro.

El primer paso en el análisis de series temporales consiste en realizar un análisis descriptivo del comportamiento pasado de la misma. El análisis de series temporales ha de iniciarse con una **representación gráfica de la misma**, de forma que en el eje de abscisas se represente el tiempo y en el de ordenadas, la serie observada. De esta forma, se obtendrá una serie de puntos que, al unirlos, ofrecen una visión del patrón de comportamiento de la serie de la que se pueden sacar unas primeras conclusiones de la evolución histórica de la misma.

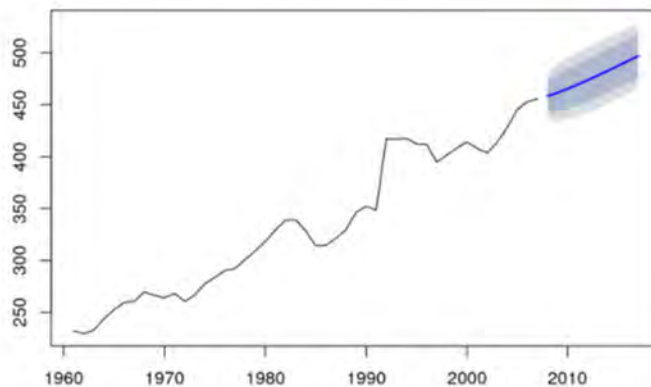
El análisis clásico de las series temporales se basa en descomponer los valores que toma la variable de observada en componentes básicos. A veces las series temporales pueden incluir tendencias, ciclos y estacionalidad. Al elegir un método de pronóstico, primero necesitaremos identificar los patrones de series de tiempo en los datos y luego elegir un método que sea capaz de capturar los patrones correctamente.

El propósito de la descomposición es aislar los distintos componentes para realizar el análisis o pronósticos sin la **influencia del ruido** o la **estacionalidad**.

## Componentes de una serie temporal

### Tendencia (T) (Trend)

Es una componente de la serie que refleja su evolución a largo plazo.



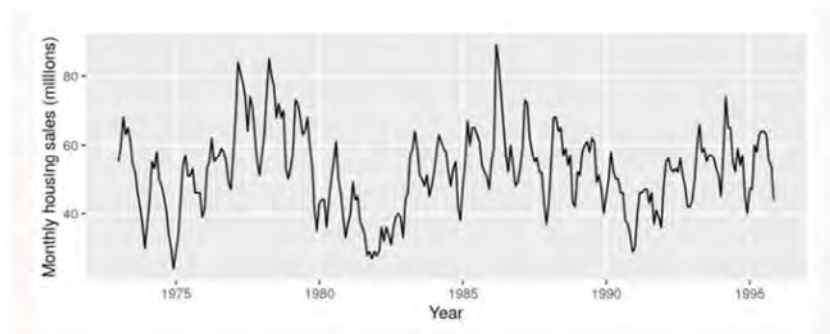
Fuente: Hyndman, R.J. *Piecewise linear trends*,  
<https://robjhyndman.com/hyndsight/piecewise-linear-trends/>

### Componente cíclica ( C ) (Cyclic)

Es una componente de la serie que recoge las oscilaciones periódicas de amplitud superior a un año y solamente es caracterizable cuando se dispone de una larga serie histórica.

Es importante diferenciar un comportamiento cíclico con el comportamiento estacional. Si la frecuencia no cambia y está asociada con algún aspecto del calendario, entonces el patrón es estacional. En general, la duración media de los ciclos es más larga que la duración de un patrón estacional, y las magnitudes de los ciclos tienden a ser más variables que las magnitudes de los patrones estacionales.

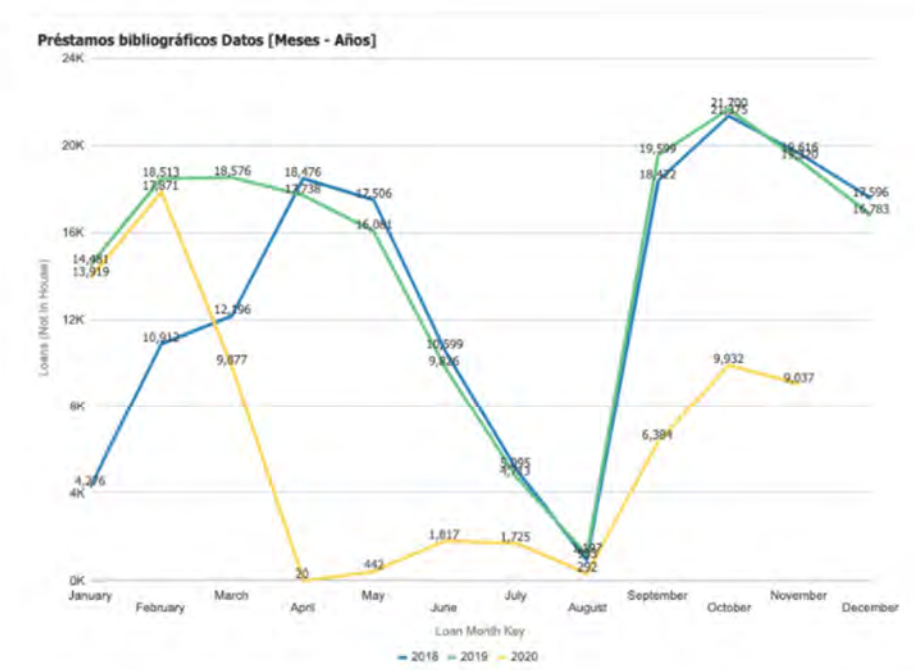
El ejemplo muestra un modelo estacional y cíclico.



Fuente: Hyndman, R.J. Cyclic and seasonal time series. <https://robjhyndman.com/hyndsight/cyclitics/>

### Componente estacional (S) (Seasonal)

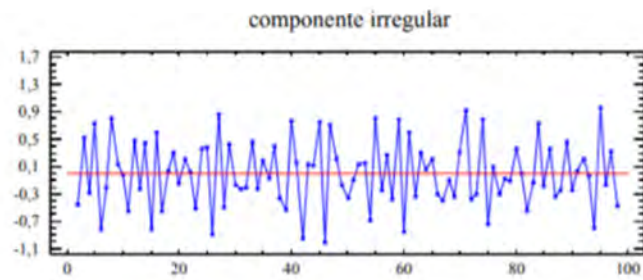
Es un movimiento periódico de corto plazo. Se trata de un componente causal debido a la influencia de ciertos fenómenos que se repiten de manera periódica en un año. Recoge las oscilaciones que se producen en esos periodos de repetición.





### Componente aleatorio o irregular (I) (Irregular)

Es una componente irregular e impredecible de las series temporales. Es por definición irregular y describe las influencias aleatorias de la misma y actúa en cualquier serie temporal en mayor o menor medida.



## Clasificación de las observaciones de las series temporales

### a) Procesos estacionarios.

Las series que toman valores estables alrededor de un nivel constante (no depende del tiempo), sin mostrar una tendencia a crecer o decrecer a largo plazo se denominan '**Procesos estacionarios**'. Por ejemplo, la cantidad de lluvias anuales en una región, o la proporción de nacimientos que corresponden a niñas.

Una **serie de tiempo estacionaria** es aquella cuyas propiedades no dependen del momento en que se observa la serie. Así, las series de tiempo con **tendencias**, o con **estacionalidad**, **no son estacionarias** - la tendencia y la estacionalidad afectarán el valor de la serie en tiempos diferentes.

El concepto de '**series estacionarias**' deriva del concepto de física que significa 'que se reproduce de forma idéntica a lo largo del tiempo'. Son series temporales que se han comportado de manera constante en el pasado y se puede suponer que se seguirán comportando de la misma forma en el futuro.

Los modelos de predicción de series temporales están diseñados para procesos estacionarios. Si las características del proceso cambian a lo largo del tiempo, resultará difícil representar la serie para intervalos de tiempo pasados y futuros mediante un modelo lineal sencillo. Sin embargo, gran parte de las series temporales suelen tener tendencias crecientes o decrecientes, y variabilidad no constante. Esta limitación no es tan importante porque, en la práctica, se pueden transformar las series no estacionarias en otras que sí lo son.

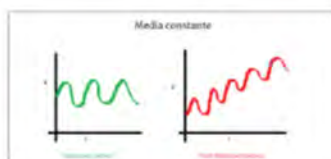
La mayoría de los modelos que intentan predecir el comportamiento de las series temporales lo hacen bajo los supuestos de que los procesos estacionarios sean estables a lo largo del tiempo y con un doble requisito:

- Una **media** aproximadamente constante en el tiempo.
- Una **varianza** o dispersión constante en el tiempo.

- **Media constante**

La serie de la izquierda tiene una media constante, mientras que la figura de la derecha muestra una tendencia y su media se incrementa con el paso del tiempo.

En términos gráficos, una media constante supone la no existencia de tendencia. Aparece dibujada como una línea es paralela a las abscisas.

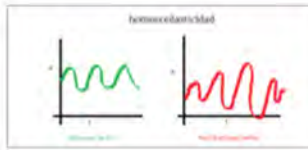


Fuente: [www.seanabu.com](http://www.seanabu.com)

- **Homoscedasticidad**

**Homoscedasticidad** u homogeneidad de **las varianzas**.

La serie de la izquierda es estacionaria. En términos gráficos, corresponde a un gráfico en el que las oscilaciones alrededor de la media son similares.

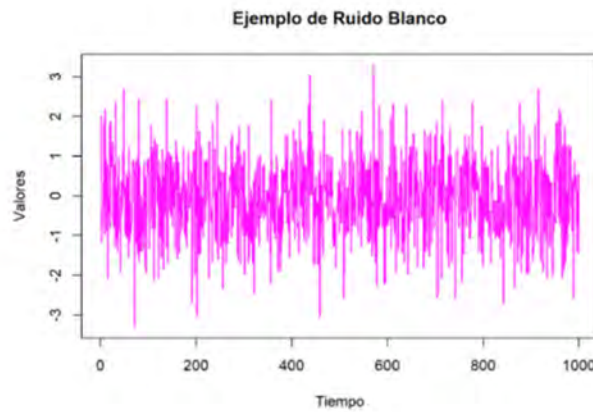


Fuente: [www.seanabu.com](http://www.seanabu.com)

**b) Variable error o ruido blanco**

Es un proceso estacionario particular.

El ruido blanco está formado por variables aleatorias de perturbaciones no observables ocurridas en el pasado con distribución normal, con media cero y varianza constante.



### c) Procesos estacionales

Corresponden a los movimientos de una variable que ocurre durante una frecuencia homogénea de tiempo, cuya periodicidad puede ser diaria, semanal, mensual, trimestral. Este proceso se caracteriza por aparecer en un periodo y desvanecerse en el siguiente. Los movimientos estacionales responden a un patrón. Suelen ocurrir con una estructura de dependencia estacional nuevamente después de un lapso equivalente de tiempo.

Los siguientes gráficos pueden ilustrar las diferencias entre los procesos estacionales y los procesos con estacionariedad:

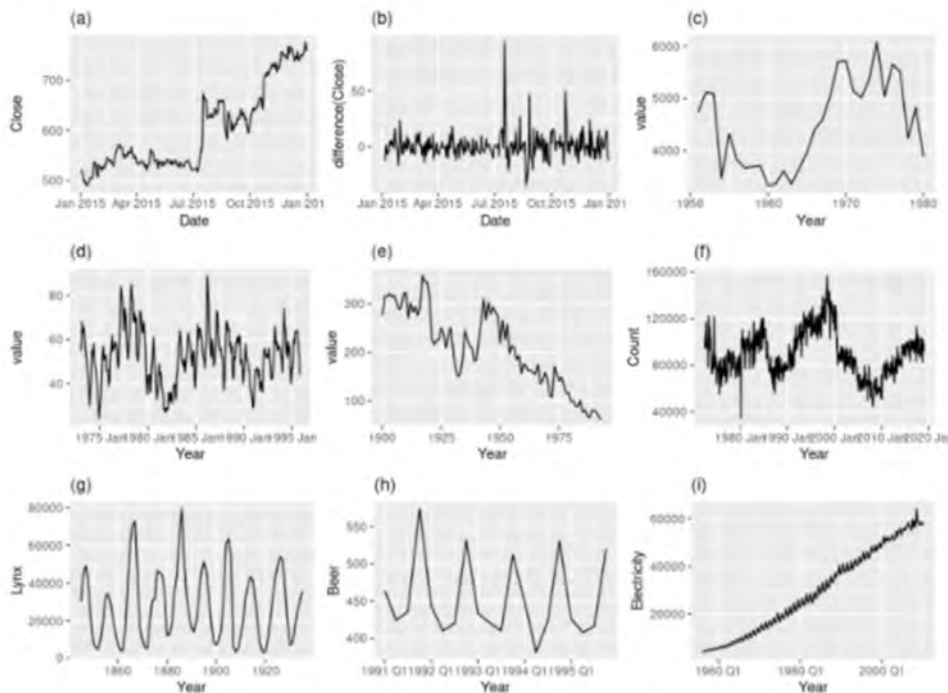


Figure 9.1: Which of these series are stationary? (a) Google closing stock price in 2015; (b) Daily change in the Google stock price in 2015; (c) Annual number of strikes in the US; (d) Monthly sales of new one-family houses sold in the US; (e) Annual price of a dozen eggs in the US (constant dollars); (f) Monthly total of pigs slaughtered in Victoria, Australia; (g) Annual total of Canadian Lynx furs traded by the Hudson Bay Company; (h) Quarterly Australian beer production; (i) Monthly Australian electricity production.

Fuente: Otexts:9.10 ARIMA vs ETS | Forecasting: Principles and Practice

<https://otexts.com/fpp3/stationarity.html>

Las series (b) y (g) son series estacionarias. Las series (d), (h), e ('i') son estacionales. Las tendencias y los niveles cambiantes excluyen las series (a), (c), (e), (f) e (i). El aumento de la varianza también descarta (i). Aparentemente, los fuertes ciclos de la serie (g) pueden parecer que la hacen no estacionaria. Pero estos ciclos son aperiódicos: se producen cuando la población de linces se vuelve demasiado grande para el alimento disponible, de modo que dejan de reproducirse y la población cae a números bajos, luego la regeneración de sus fuentes de alimentos permite que la población vuelva a crecer. A largo plazo, la sincronización de estos ciclos no es predecible. Por tanto, la serie es estacionaria.



## Modelos ARIMA

Los modelos ARIMA son técnicas para el análisis de series temporales, fueron desarrollados por Bob y Jenkins (2). La primera cuestión que se plantea es de orden terminológico, qué se entiende por modelos ARIMA y cuál es su finalidad.

Los modelos ARIMA (*AutoRegresive Integrated Moving Average*), cuyo nombre deriva de de los componentes **AR** (Autorregresivo), **I** (Integrado) y **MA** (Medias Móviles), *Modelos Autorregresivos Integrados de Medias Móviles*, son modelos estadísticos paramétricos que tratan de explicar el presente de una serie temporal por lo que ha sucedido en el pasado. Es decir, cada observación en un momento dado es modelada en función de valores anteriores.

El modelo ARIMA permite describir un valor en función de los datos anteriores y errores debidos al azar, además de modelizar y predecir series estacionarias y no estacionarias a las que se les ha eliminado la tendencia y la estacionalidad. Bob y Jenkins recomiendan como mínimo 50 observaciones en la serie temporal.

El objetivo de los modelos ARIMA es predecir una serie de datos de forma no determinista, ya que contiene un componente aleatorio (error o ruido blanco).

## Modelos autorregresivos (AR)

**i** Los modelos autorregresivos **AR** tratan de explicar una variable por lo sucedido en el período precedente. Es decir se realiza una regresión consigo misma desfasada un periodo o retardo. El término autorregresión indica que es una regresión de la variable contra sí misma.

En un modelo de autorregresión, se pronostica la **variable dependiente** utilizando una combinación lineal de valores anteriores de la variable.

- **Coefficientes de autocorrelación**

Es el instrumento fundamental para analizar las series temporales en términos de observaciones. El coeficiente de autocorrelación mide la correlación, es decir, el grado de asociación lineal que existe entre observaciones separadas por ***k tiempo***. Esta función ayuda a conocer la dependencia que tienen los datos en un período determinado con los mismos 'k periodos' anteriores.

Los coeficientes de autocorrelación ayudarán a construir el modelo apropiado para los datos.

El **coeficiente de autocorrelación** del mismo modo que el **coeficiente de correlación** toma los valores entre -1 y +1, siendo '0' la ausencia de autocorrelación y -1 y +1 la autocorrelación perfecta.



## Modelo de medias móviles (MA)

- Una alternativa para predecir una variable en base a sus valores pasados es hacerlo a través de una corrección de los 'errores observados' o 'valores del ruido blanco' de instantes precedentes, modelo denominado de 'medias móviles (MA)'.

Las medias móviles son **procesos estacionarios**. Se considera que el valor de las series temporales **MA** tienen una *media constante ( $\mu$ )* alrededor de la cual se mueve la variable. Además, se supone que la media ( $\mu$ ) en el presente está ocasionada por perturbaciones no observables ocurridas en el pasado.

En el modelo de medias móviles a todos los valores históricos observados se les dota de la misma importancia. Desde el punto más distante hasta el instante actual estos pesos se distribuyen de forma exponencial, de modo que el peso es mayor cuanto más cerca se sitúa el dato del instante actual. Los métodos basados en medias móviles son una manera simple de suavizar el histórico de datos para obtener una predicción *cuando no hay tendencia marcada ni estacionalidad*.

Una manera de modificar la influencia de los datos pasados en la media como predicción es especificar *cuántas observaciones pasadas serán incluidas en el cálculo de la media*. Se utiliza el término 'medias móviles' para describir este procedimiento puesto que conforme vamos disponiendo de nuevas observaciones se pueden calcular nuevas medias incluyendo la nueva observación y desechando la más antigua.

**La predicción de media siempre incluye el mismo número de observaciones** y éstas además, son las más recientes. Este método no es capaz de tratar tendencia o estacionalidad.

## Modelo Integrado (I)

- Modelo Integrado**, son procesos no estacionarios que ocurren en series temporales en bruto, que necesitan convertirse en estacionarias mediante diferencias en la variable.

Por ejemplo, en un modelo de predicción de ventas que intenta pronosticar el cambio en las ventas de mañana (es decir, las ventas de mañana menos las ventas de hoy) en lugar de solo las ventas de mañana. La razón es porque muchas series de tiempo muestran una tendencia, lo que hace que los valores brutos no sean estacionarios. Para que la serie sea más estacionaria se hace la diferencia de la variable.

## Modelos ARMA

Es un modelo mixto con componente autorregresiva (AR) y con componente de medias móviles (MA). Los modelos ARMA son estacionarios.

## Condiciones del modelo ARIMA

Los datos deben ser **estacionarios y univariados**. Los procesos ARIMA están diseñados para modelar datos de carácter estacionario. La metodología se hace depender a dicha variable tan sólo de su propio pasado y un conjunto de perturbaciones aleatorias.

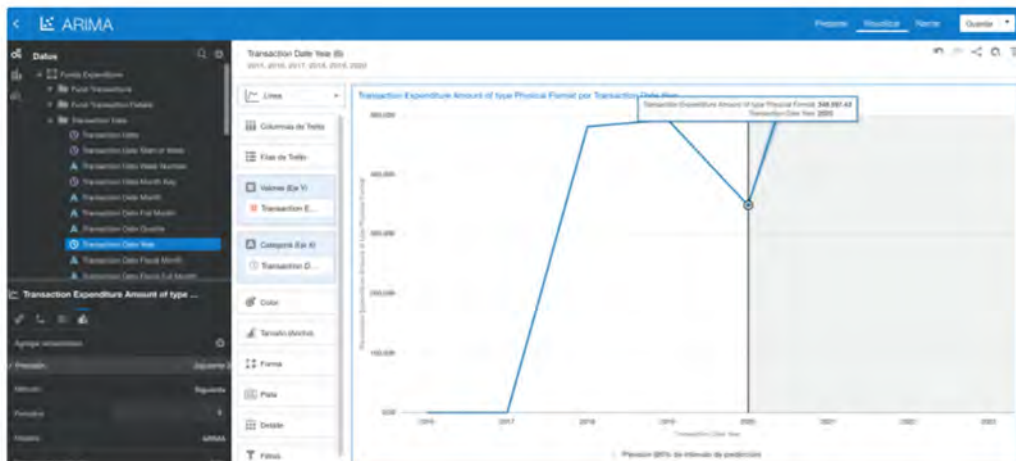
Dado que los procesos ARIMA están diseñados para modelar datos de carácter estacionario, lo primero que debemos hacer es efectuar un análisis de la estacionariedad de los datos.

Si la serie temporal que se está analizando *no es estacionaria* se deberán aplicar las transformaciones adecuadas con objeto de convertirla **en estacionaria**. La estacionariedad exige que se cumpla un doble requisito en la práctica:

- (a) Una media constante en el tiempo
- (b) Una varianza constante

Si presenta un componente 'estacional', hay que diferenciar si este componente estacional se produce de forma regular o cíclica, en este caso se trataría de **modelos SARIMA**, o si no se produce de forma no regular, **modelos ARIMA**.

Se muestra el modelo ARIMA de una serie temporal sobre inversión en recursos en papel (2015-2020), con una proyección con tendencia positiva en 2021.



## Modelos SARIMA o ARIMA con componente estacional

Los modelos ARIMA también son capaces de modelar una amplia gama de datos estacionales.

Un modelo **SARIMA** (*Seasonal Autoregressive Integrated Moving Average Model*), Modelos Autorregresivos Estacionales Integrados de Medias Móviles, se forma al incluir términos estacionales en los modelos ARIMA.

Cuando se trabaja con datos de periodicidad inferior al año (trimestrales, mensuales, o diarios), suele darse el factor estacional.

El siguiente gráfico muestra un modelo ARIMA estacional sobre los préstamos bibliográficos mensuales (Not in house) de los años 2015-2020 y la tendencia futura.

### Nota:

La herramienta 'Visualización de datos' al seleccionar 'Estacional ARIMA' proyecta de las series temporales con componente estacional a partir del criterio 'Date Year'.



### Referencias:

- (1) García Díaz, J. C. (2016). *Predicción en el dominio del tiempo*. Valencia: Editorial de la Universidad Politécnica de Valencia.
- (2) Box, G. E. P., and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association* 70: 70-79.
- (3) Box, George E. P, Jenkins, Gwilym M, & Reinsel, Gregory C. (2008). *Time Series Analysis* (Wiley series in probability and statistics). Somerset: John Wiley & Sons, Incorporated.
- (4) Otexts: 9.10 ARIMA vs ETS <https://otexts.com/fpp3/arima-ets.html>



Ejemplo:

AÑO	10.2014.001	11.100.000 €
2017	10.214.934	11.691.090 €
2018	9.535.958	10.630.954 €
2019	9.047.172	10.186.510 €

Tabla 1. Compra de monografías y préstamos generados

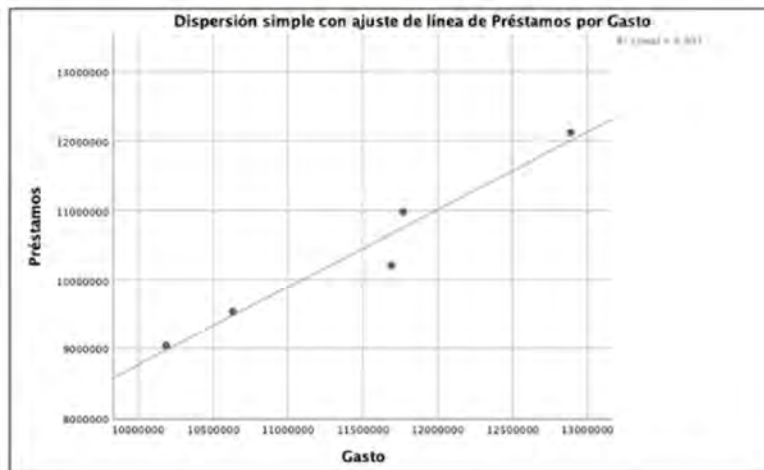


Fig. 1. Diagrama de dispersión de los préstamos generados en relación con el gasto en monografías

### Modelo de regresión lineal simple

Dos variables 'x' e 'y' están relacionadas según una **línea recta** cuando presentan una asociación lineal. Esto significa que al tomar una diferentes valores, los valores medios de las otra se modifican.

Una **recta inclinada** indicaría que las variables presentan una asociación lineal, ya que al tomar una de ellas diferentes valores, los valores medios de la otra se modifican. En este ejemplo, la línea inclinada (Fig. 1) indica que el gasto y los préstamos están relacionados linealmente.

Si las variables siguen una **recta horizontal** indicaría que las variables no están linealmente relacionadas, ya que al tomar una de ellas diferentes valores, los valores medios de la otra no cambian. En este ejemplo, la recta horizontal indicaría que el número de préstamos no varían aunque el gasto invertido en la compra de monografías variase.

El lugar geométrico de las medias de las distribuciones ligadas se denomina **línea de regresión**.

La recta de regresión  $R_{yx}$  inclinada cuya pendiente  $>0$ , sugiere una **asociación lineal positiva**, mientras que si la pendiente de la recta  $R_{yx}$  inclinada  $<0$ , sugiere una asociación lineal negativa (Véase Fig.3).

Dos variables están **asociadas positivamente** cuando los valores superiores a la media de una de ellas tienden a ir acompañados de valores también situados por encima de la media de la otra variable, y cuando los valores inferiores a la media también tienden a ocurrir conjuntamente.

Cuando la recta de regresión  $R_{yx}$  inclinada (pendiente inferior a 0) las variables están **asociadas negativamente**. Los valores superiores a la media de una de ellas tienden a ir acompañados de valores inferiores a la media de la otra variable, y viceversa.

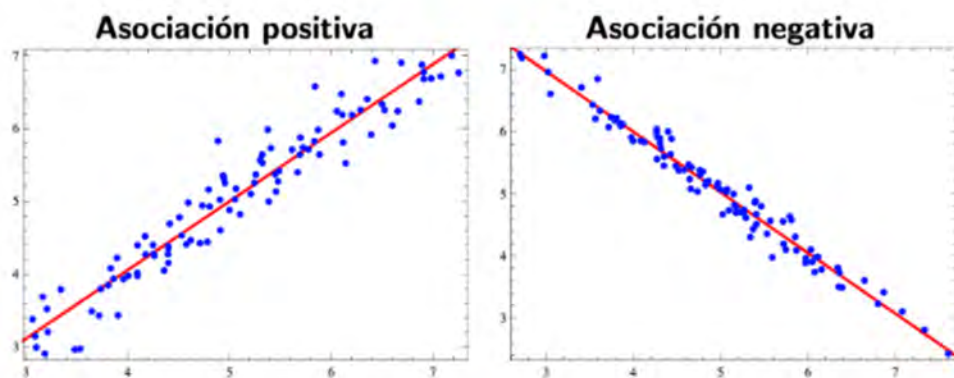


Fig. 2. Asociación entre variables

### Planteamiento del modelo de regresión

Se parte de la hipótesis que las dos variables del estudio son independientes, y de una hipótesis alternativa en la que se supone que las dos variables están relacionadas.

Hipótesis de partida. Las dos variables son independientes ( $H_0$ ). Sin significación:  $p\text{-valor} > 0,05$

Hipótesis alternativa. Las dos variables del estudio están relacionadas ( $H_a$ ). Con significación:  $p\text{-valor} < 0,05$

Para determinar si las variables están relacionadas o no, conocer la dirección y fuerza de la relación se calculará el coeficiente de correlación de Pearson y se comprobará la significación.

La correlación de Pearson es un estadístico paramétrico que mide si los puntos tienen tendencia a disponerse en línea recta. El objetivo de la correlación de Pearson es indicar el grado de asociación entre dos variables, de forma que al aumentar el valor de una variable aumenta la otra y viceversa. Este concepto de variación conjunta se llama covarianza. El coeficiente de Pearson se calcula dividiendo la covarianza entre las desviaciones típicas.

$$r = \frac{S_{XY}}{S_X S_Y}$$

Correlaciones			
		Préstamos	Gasto
Correlación de Pearson:	Préstamos	1,000	,976
	Gasto	,976	1,000
Sig. (unilateral):	Préstamos	.	,002
	Gasto	,002	.
N	Préstamos	5	5
	Gasto	5	5

Tabla 1: Correlaciones. Salida SPSS (ejemplo, préstamos-gasto)

$r=0,976$  indica que existe una relación positiva, fuerte y significativa entre las variables 'préstamo' y 'gasto', con un  $p\text{-valor}=0,002$ , luego existe significación estadística.

La correlación explica la fuerza de la relación lineal de las variables. El siguiente paso consistirá en buscar **el modelo de regresión** que mejor explique la variación de 'y' cuando varía 'x', es decir el hecho de que cuando cambie 'x' (gasto) arrastre consigo a 'y' (préstamos).



### Ecuación de la recta de regresión

Decimos que dos variables 'x' e 'y' están relacionadas **según una línea recta** cuando sus valores satisfacen la ecuación:

$$y = a + bx$$

Esta ecuación que describe la relación entre las dos variables se denomina '**Modelo de regresión**', donde 'a' y 'b' son constantes:

- **a: intercepto**, o punto de corte de la recta en el eje de ordenadas.
- **b: geoméricamente se refiere a la pendiente o inclinación de la recta**. Es decir, indica cuánto cabe esperar que cambie la respuesta de la variable dependiente 'y' por cada incremento de 'x'.
- '**x**' e '**y**': son las variables.

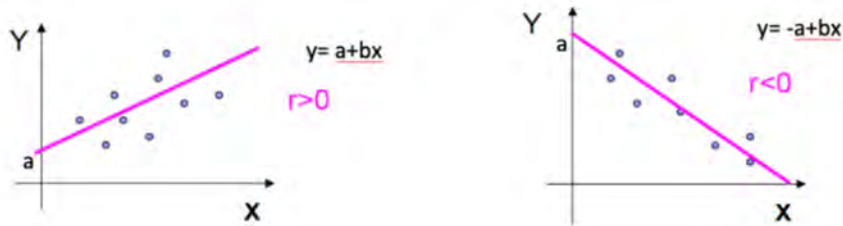


Fig. 3: Ecuación que describe la recta de regresión

### Cálculo de la recta de regresión

Para hallar la recta que mejor se ajusta a la nube de puntos la construcción de los modelos se basa en la '**Teoría de los Mínimos Cuadrados**'. Se ilustra el ajuste de la recta en la Fig. 4.

Se traza una recta que mejor describa la nube de puntos a partir de la ecuación ( $y = ax + b$ ). Esta recta contiene los **valores estimados ( $y_i^*$ )**. Por tanto, en la gráfica nos encontramos con el valor **verdadero u observado ( $y_i$ )**, y con el **valor estimado ( $y_i^*$ )**.

Esto significa según el ejemplo anterior (gasto-préstamos) que la Fig. 4 describiría un **valor observado** para la ' $x_i$ ' (gasto) y un **valor observado** para la ' $y_i$ ' (préstamos), representados en el punto ( $x_i, y_i$ ). Sin embargo, el modelo matemático proporcionaría un valor estimado ( $x_i, y_i^*$ ).

Obviamente, interesará minimizar la distancia entre el valor observado y el valor estimado por el modelo. De ahí que esta teoría se denomine de '**Mínimos cuadrados**' porque trata de minimizar la distancia entre los puntos (valores observados) y la recta ajustada [ $y_i^* = a + bx_i$ ] (valores estimados) al cuadrado, porque esta diferencia '**residuo**' o '**error de predicción**' se eleva al cuadrado (Véase fig. 5).

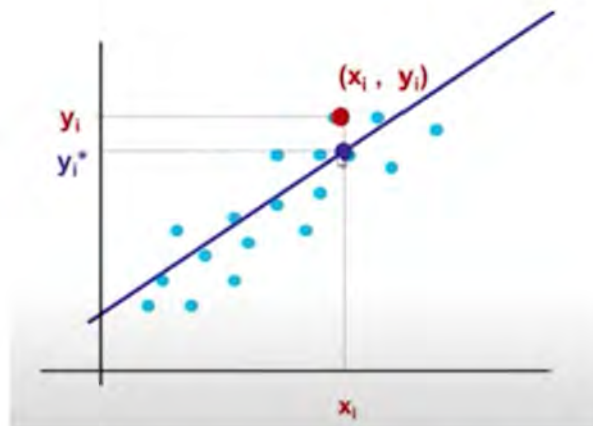


Fig. 4. Ajuste de una recta a una nube de puntos. Fuente: Universidad de Salamanca

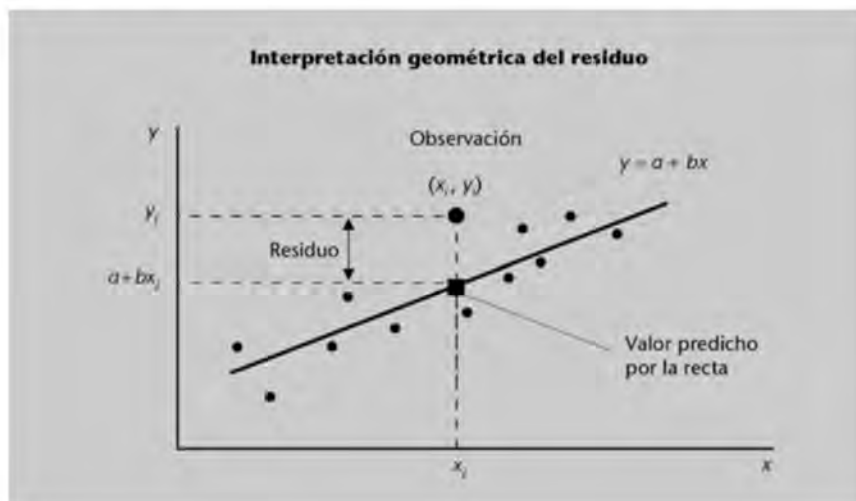


Fig. 5. Interpretación geométrica del residuo. Fuente: Instituto Aragonés de Ciencias de la Salud

La ecuación es la suma de los residuales al cuadrado.

$$D = \sum (y_i - y_i^*)^2 = \sum (y_i - (a + bx_i))^2 = \sum (y_i - a - bx_i)^2$$

Para cada observación, hay una diferencia entre el punto y la recta llamada '**residual**', que representa el error de predicción.

### Cálculo del coeficiente de regresión 'b' y 'a'

Para hallar 'b', denominado **coeficiente de regresión** de Y sobre X, se divide la covarianza por la varianza de la variable independiente 'x'. Para calcular 'a' se halla la media de la variable 'y', menos el coeficiente de regresión 'b', por la media de los valores de 'x'.

$$b = \frac{S_{xy}}{S_x^2} \quad a = \bar{y} - b\bar{x}$$

La covarianza  $S_{xy}$ , es una medida de variabilidad conjunta de dos variables respecto a sus medias. La covarianza nos indica si la relación entre dos variables es:

- Directa:  $S_{xy} > 0$
- Inversa:  $S_{xy} < 0$
- Incorreladas:  $= 0$

La varianza  $S_x^2$ , es una medida de dispersión que se utiliza para calcular la variabilidad de un conjunto respecto a la media de los mismos. La varianza es **la desviación típica al cuadrado**.

La varianza, por definición siempre es positiva, ya que es una suma de valores al cuadrado. Por tanto, el símbolo que tenga la covarianza es el que tendrá el coeficiente de regresión 'b'.

### Evaluación de la bondad de ajuste del modelo

La bondad de ajuste del modelo se evalúa por el **coeficiente de determinación  $R^2$** .

El coeficiente de determinación  $R^2$  se obtiene a partir de la razón entre la variación explicada y la variación total:

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}} = \frac{\sum (y_{est} - \bar{y})^2}{\sum (y - \bar{y})^2}$$

$R^2$ : el coeficiente de determinación toma los valores entre 0 [ajuste nulo] y 1 [ajuste perfecto]. Cuanto más se aproxima  $R^2$  a '1' mayor poder explicativo tendrá el modelo. En la práctica, este coeficiente se indica en porcentaje.

### Interpretación del ejemplo (gasto-préstamos)

Coeficientes <sup>a</sup>							
Modelo	Coeficientes no estandarizados		Coeficientes estandarizados		Sig.	95,0% intervalo de confianza para B	
	B	Desv. Error	Beta	t		Límite inferior	Límite superior
1	(Constante)	-2426931,3		-1,477	,236	-7657912,3	2804049,77
	Gasto	1,120	,143	,976	,004	,664	1,576

a. Variable dependiente: Préstamos

La constante 'a' = -2426931,3 .

La constante 'a' es el intercepto, el punto de corte de la recta en el eje de las ordenadas (préstamos).

'a' = -2426931,3 indicará el dato estimado de préstamos cuando 'x'=0.

**Coefficiente de regresión 'b'**= 1,120. Es el número de préstamos / euro gastado. Este dato indica que por cada unidad que aumente el gasto, el préstamo aumenta en 1,120.

**P-valor**= 0,004. Significación estadística. Es decir, el resultado no es un producto del azar del muestreo e implica que la variable 'x' tiene capacidad para describir las variaciones de la variable 'y'. P-valor=0,236 no es significativo. Este dato corresponde al intercepto. Se puede eliminar el intercepto volviendo a ejecutar el modelo. El nuevo modelo cambiará la pendiente.

La ecuación de la recta de regresión es:  $Y = -2426931,3 + 1,120x$

Resumen del modelo <sup>b</sup>									
Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Cambio en R cuadrado	Cambio en F	Estadísticos de cambio		Sig. Cambio en F
1	,976 <sup>a</sup>	,953	,938	303681,706	,953	61,129	gl1	gl2	
							1	3	,004
a. Predictores: (Constante), Gasto									
b. Variable dependiente: Préstamos									

**R<sup>2</sup> = 0,953. Coeficiente de determinación.** Este resultado indicaría que el el modelo explica el 95,3% de la variación del préstamo en relación con la inversión en la compra de libros. El modelo es muy bueno porque los puntos están muy cerca de la recta.

**r = 0,976. Coeficiente de correlación de Pearson.** Los valores de correlación pueden tomar un rango de [-1 a +1]. Cuando *r* esté próximo a +1 ó -1 la correlación indicará que las variables están intensamente relacionadas. Un valor próximo a '0' indicará que apenas hay asociación entre las dos variables.

Se ha demostrado que también se puede calcular el coeficiente de determinación **R<sup>2</sup>** elevando al cuadrado el índice de correlación de Pearson **r<sup>2</sup>**. De forma que:

$$R^2 = 0,953, \quad r^2 = 0,953$$

**R cuadrado ajustado:** 0,938. Este dato se emplea en la regresión lineal múltiple.

**Error estándar de estimación (SE).** Es la desviación estándar de la distribución del muestreo.

**gl1 y gl2:** grados de libertad, o variables explicativas que tiene el modelo. El grado de libertad g1 es el número de variables -1, Por tanto, gl1 tiene un grado de libertad.

**Cambio en F (Anova)**= 61,129. El estadístico F del análisis de la varianza (Anova) contrasta la la igualdad de las muestras. El estadístico F permite contrastar la hipótesis nula ( $H_0$ ) de la pendiente de la muestra. Cuando  $F=0$ , las medias muestrales son idénticas. Cuando el valor  $F=1$  no existe significación y se acepta  $H_0$ . Los valores F grandes constituyen una buena evidencia en contra de la hipótesis nula.

**Significación cambio en F (Anova):** p-valor 0,004. significación estadística.



### Condiciones de la regresión lineal. Análisis de residuos

Una vez construido el modelo de regresión se comprueban a posteriori los supuestos. Se utiliza el análisis de residuos para realizar su comprobación. Los residuos o errores se definen como la diferencia entre el valor observado y el valor estimado (Véase Fig. 5). Supuestos:

-Linealidad: la relación de ambas variables debe ser lineal. Este supuesto se puede comprobar utilizando la gráfica de ambas variables y superponiendo la recta del modelo generado por regresión.

-Distribución normal de los residuos. Los residuos se deben distribuir de forma normal, con media igual a '0'. Los valores extremos suelen ser una causa por la que se viola la condición de normalidad.

**Estadísticas de residuos<sup>a</sup>**

	Mínimo	Máximo	Media	Desv. Desviación	N
Valor pronosticado	8984029,00	12009879,0	10380378,2	1187167,51	5
Residuo	-454529,88	224638,172	,000	262996,072	5
Desv. Valor pronosticado	-1,176	1,373	,000	1,000	5
Desv. Residuo	-1,497	,740	,000	,866	5

a. Variable dependiente: Préstamos

-Homocedasticidad: homogeneidad o igualdad de las varianzas. Se comprueba en una gráfica si los residuos se distribuyen de una manera constante a lo largo del eje X. Se comprueba este supuesto se puede utilizar la prueba de Levene.

-Valores atípicos. Hay que estudiar los valores atípicos o extremos ya que pueden generar falsas correlaciones u ocultar las existentes.

-Independencia. Las observaciones deben ser independientes unas de otras. Es importante tener en cuenta esta condición en el caso de mediciones temporales. En este caso la compra de los préstamos son independientes en el tiempo y los datos han sido medidos de forma independiente.

### Estimación y predicción

Uno de las aplicaciones más frecuentes de la regresión lineal es la estimación y predicción del valor de 'y' en función de 'x'. Por ejemplo, si se quiere estimar el número de préstamos para una inversión de 8 millones €.

$$Y = -2.426.931 + 1,120 * 8.000.000 = 6.533.069 \text{ préstamos.}$$

El modelo de predicción está diseñado para estos valores mínimos y máximos estudiados. Extender los datos puede crear incertidumbre porque no se puede saber cómo se van a comportar los datos. La utilidad del modelo dependerá del conocimiento del campo de aplicación y conocimiento del comportamiento de los datos.